

Institute of Mathematics
Polish Academy of Sciences

Michał Józwikowski

Optimal control theory on almost Lie algebroids

PhD dissertation

Supervisor:
prof. dr hab. Janusz Grabowski

Institute of Mathematics
Polish Academy of Sciences

Warsaw, June 2011

Acknowledgements

This work would not be completed without the help of my supervisor—professor Janusz Grabowski. I would like to thank him for many inspiring discussions, useful suggestions (including the choice of this topic) and, especially, for his patience and tolerance.

I would also like to express my gratitude to the whole scientific and non-scientific staff of the Institute of Mathematics where I spent the last four years.

Finally, I thank my Parents and Olga for their constant support.

Contents

Acknowledgements	i
Contents	ii
1 Introduction	1
An overview	1
Optimal control theory	1
Algebroids and their application to mechanics	2
Reduction in optimal control theory	3
The main result	4
Discussion of the main result	7
Organisation of the manuscript	9
2 Almost Lie algebroids	11
Differentiable manifolds and vector bundles	11
Almost Lie algebroids	11
Hamiltonian vector fields and tangent lifts	12
Cartan Calculus	13
Morphisms	14
Admissible paths	14
The product of skew-algebroids	15
3 Homotopies of admissible paths	17
3.1 The E -homotopy and its interpretation	17
Uniqueness of E -homotopies	19
The E -homotopy via Stokes theorem	19
E -homotopies on an integrable algebroid	21
E -homotopy classes	24
The interpretation of E -homotopy classes	25
3.2 Fundamental properties of E -homotopies	26
E -homotopies as families of E -paths	26
Relation between E -homotopies with and without fixed end-points, reparametrisation	28
4 Optimal control problems	31
Control systems on AL algebroids	31

Optimal control problems on AL algebroids	32
Interpretation of the algebroid OCPs	33
OCPs in terms of the product algebroid $E \times T\mathbb{R}$	35
The algebroid homotopy associated with a control system	36
5 The Pontryagin Maximum Principle	39
Formulation on an AL algebroid E	39
Alternative formulation	40
6 Discussion of the main result	43
6.1 Non-autonomous versions of the PMP	43
6.2 The known results	45
The PMP on Lie groups and principal bundles	45
An application—the falling cat problem	46
Applications to variational problems	47
6.3 Other examples	49
Two-point time OCP on $\mathfrak{so}(3)$	49
An application to a nonholonomic system	50
7 Needle variations	53
Needle variation of controls and trajectories	53
Needle variations and \mathbf{A} -homotopy classes	54
The set of infinitesimal variations K_τ^u	59
8 Technical lemmas	63
Local coordinates	63
First lemma	64
Second lemma	70
9 The proof of the PMP	73
9.1 The proof of Theorem 5.3	73
The geometry of the cone K_τ^u	73
The construction of $\xi(t)$ and the “maximum principle”	75
The condition $\mathbf{H}(\mathbf{x}(t), \xi(t), u(t)) = 0$	76
Extending the “maximum principle” to $[t_0, t_1]$	78
9.2 The proof of Theorem 5.4	79
A Differential geometry	83
A.1 Lie groupoids	83
Lie groupoids	83
A Lie algebroid of a Lie groupoid	84
Lie theory	86
A.2 The Atiyah algebroid	87
Invariant vector fields	88
The Atiyah sequence	88

Local description	89
The linear Poisson structure	90
B Analysis	91
B.1 Mesuaralbe maps, regular points	91
Basic facts	91
Uniform regularity	92
B.2 Ordinary differential equations	98
Carathéodory solutions	98
Parameter dependence	99
Gronwall Inequality	100
B.3 E -homotopy type equations	100
E -homotopy type equations	100
Existence of the trace and properties of WT-solutions	101
C Control theory	105
D Geometry and Topology	107
D.1 Separation of convex cones	107
D.2 Simple topological lemmas	109
Bibliography	111
Index	117

Introduction

An overview

This work is rooted in two important areas of mathematics. One of them is the optimal control theory with its central theorem, the celebrated Pontryagin maximum principle (PMP). The second is the theory of Lie algebroids and, in particular, its applications to geometric mechanics.

Note that the PMP may be regarded as an extension of the calculus of variations to a much bigger class of problems. On the other hand, the language of Lie algebroid theory has proved to be a very fruitful tool in mechanics and variational calculus allowing, for example, to treat standard systems and systems reduced by inner symmetries in a unified way, and to have a deeper insight into the nature of the Lagrange and Hamilton formalisms. Therefore, we may think of the PMP and the algebroidal formulation of the Lagrange formalism as two different extension of the standard calculus of variations—the first by generalising the class of problems, the second by generalising the geometric context.

Consequently, it is a natural idea to unify these two generalisations and formulate the PMP in the language of algebroids. Some attempts in this direction has already been made ([Cortés & Martínez, 2004, Martínez, 2004, Grabowska & Grabowski, 2008]), yet so far there is no satisfactory solution. In this work we tried to give a full solution of the problem in a possibly general context both geometrical and technical. Namely, we formulate our extension of the PMP for optimal control problems (OCPs) on almost Lie algebroids (objects generalising Lie algebroids), we work with bounded measurable controls and absolutely continuous base trajectories, and we consider quite general boundary conditions. Note that, according to [Agrachev & Gamkrelidze, 2006], the PMP was never a subject of any substantial generalisation, apart from the technical ones discussed in the next subsection.

Optimal control theory

The PMP, proved in 1956 by L. Pontryagin and his collaborators [Pontryagin *et al.*, 1962], was an answer to the problem of finding solutions of optimal problems of new type which could not be treated with the standard variational methods. Roughly speaking, we are interested in minimising the standard action functional of the calculus of variation (on a manifold M), but we restrict our attention to trajectories $x(t) \in M$ whose velocities can

be controlled, i.e., they belong to a given subset of TM . The PMP (Theorem C.2) expresses the necessary conditions for optimality in the language of the canonical symplectic structure of the cotangent bundle T^*M . A short account of the result and the historical discussion of its development can be found in [Agrachev & Gamkrelidze, 2006].

From its appearance, the PMP became an object of intensive studies both on theoretical and applied level. As a technique it is used in a wide range of disciplines which include engineering, aerospace, robotics, medicine, economics, and other (see the references in [Barbero-Liñán & Muñoz-Lecanda, 2009] for more details), and as such is used for solving concrete practical problems.

The theoretical development concentrated in several directions. One of them, initiated by Clarke in the 70s, was devoted to relaxing the assumptions under which the result holds. This research, undertaken mostly by Clarke, Ioffe, Loewen, Mordukhovich, Rockafellar, and Vinter, used the tools provided by the non-smooth analysis and led to generalisations of the PMP among which the most important is [Clarke, 1976]. The monograph [Clarke, 2005b] discusses this topic in details (see also [Clarke, 2005a] for a brief account of the most important results).

In the 60s there was a search for a simple proof of the PMP. The original argument of Boltyanskii ([Pontryagin *et al.*, 1962]) is long, and some people believed that a shorter reasoning based on variational methods can be found. Such a proof does not exist so far, even though some simpler versions of the PMP can be proven quite elementary. Essentially, there are two kinds of proofs of the PMP. The first is the original one which uses needle variations—a tool developed by Boltyanskii especially for this purpose. The other argument was given by [Gamkrelidze *et al.*, 1978] and is based on the concept of generalised controls. The research in this area concentrated mostly on translating the proof expressed originally in the language of differential equations to the language of differential geometry. Recent results [Barbero-Liñán & Muñoz-Lecanda, 2009, Agrachev & Sachkov, 2004] show a deeper understanding of the geometric origins on the PMP and connection between optimality and accessibility.

The last topic is closely related to the problem of abnormal extremals. These are the solutions of the OCPs which do not depend on the cost function but on the geometry of the considered system only. For a long time, until the discovery of counterexamples in sub-Riemannian geometry [Montgomery, 1994], people believed that such curves cannot be optimal. Since then this area became a subject of a growing interest [Agrachev & Sarychev, 1996, Agrachev & Sarychev, 1998, Bonnard & Trélat, 2001, Langerock, 2003a, Langerock, 2003b].

Algebroids and their application to mechanics

Lie algebroids were introduced by Pradines in the mid 60s as infinitesimal objects associated with Lie groupoids, per analogy to Lie algebras and Lie groups. In a series of short articles [Pradines, 1966, Pradines, 1967b, Pradines, 1967a, Pradines, 1968] he announced a very general program of developing the Lie theory for Lie groupoids. The progress was not very fast until the 80s, when Weinstein introduced the notion of a symplectic groupoid to Poisson geometry [Weinstein, 1987]. Since then Lie al-

gebroids and Lie groupoids has become objects of great significance in this field (see e.g. [Coste *et al.*, 1987, Weinstein, 1988, Courant, 1990, Weinstein & Xu, 1991, Xu, 1992, Cattaneo & Felder, 2004, Crainic & Fernandes, 2004]). The main reason of this is the fact that with every Poisson manifold P one can naturally associate a Lie algebroid structure on the cotangent bundle $T^*P \rightarrow P$ (on the other hand Lie algebroids are objects dual to linear Poisson structures). Therefore many problems of Poisson geometry can be translated into the language of Lie algebroid theory. Other applications of Lie algebroids appeared in the theory of foliations (e.g. [Pradines, 1966, Winkelkemper, 1983, Moerdijk & Crainic, 2001, Moerdijk & Mrčun, 2003]) and, for locally trivial Lie algebroids, in the theory of connections (cf. [Mackenzie, 1987]).

In all these applications the problem of integrability of Lie algebroids posted already by Pradines plays a central role. For example the existence of a symplectic realisation of a Poisson manifold P is equivalent to the integrability of the associated Lie algebroid $T^*P \rightarrow P$. The integrability problem was attacked by many authors and partial solutions for some special classes of Lie algebroids has been obtained (see [Almeida, 1980, Almeida & Kumpera, 1981, Almeida & Molino, 1985, Mackenzie, 1987, Cattaneo & Felder, 2004]) until it was completely solved by [Crainic & Fernandes, 2003]. A detailed discussion of this topic is given in Appendix A.1.

We are interested mostly in applications of the algebroid theory to mechanics which was also a Weinstein's idea [Weinstein, 1996] (see also [Libermann, 1996]). Since then the topic was studied in different contexts by many authors ([Cortés *et al.*, 2006, Cortés & Martínez, 2004, León *et al.*, 2005, Martínez, 2001a, Martínez, 2001b, Martínez, 2005, Martínez, 2007, Martínez, 2008]). It was observed a little bit later, following the approach to analytical mechanics proposed by Tulczyjew [Tulczyjew, 1974, Tulczyjew & Urbanski, 1999], that geometrical mechanics, together with the Euler-Lagrange and the Hamilton equations, constrained dynamics, etc., can be developed based on more general objects than Lie algebroids ([Grabowska *et al.*, 2006, Grabowska & Grabowski, 2008]). They were introduced in [Grabowski & Urbanski, 1999] under the name (*general*) *algebroids*. This generalisation turns out to be of practical use, as systems of mechanical type with nonholonomic constraints allow a nice geometrical description in terms of *skew-algebroids* [Grabowski *et al.*, 2009] which do not have to satisfy the Jacobi identity in general.

Reduction in optimal control theory

As a motivation, before formulating our main results, let us discuss a reduction by inner symmetries of a control or mechanical system. It is a well-known phenomena in analytical mechanics and control theory that symmetries of a system lead to reductions of its degrees of freedom. It is also well-understood that such a reduction procedure is not purely computational but is associated with a reduction of the geometrical structures hidden behind.

A typical situation considered in control theory is a control system $F : P \times U \rightarrow TP$ on a manifold P (with U being the set of control parameters) which is equivariant w.r.t. the action of a Lie group G on P and the induced action on TP . If this action is free and proper, we deal in fact with a G -invariant control system on a principal bundle $P \rightarrow P/G$.

Introducing a G -invariant cost function $L : P \times U \rightarrow \mathbb{R}$, one ends up with a G -invariant optimal control problem on a principal bundle $P \rightarrow P/G$.

There are basically two ways of obtaining optimality necessary conditions for such a problem. In the first, one takes the PMP for the unreduced system on P and performs the Poisson reduction of the associated Hamiltonian equations. For the simple case of an invariant system on a Lie group $P = G$ (see eg. [Jurdjevic, 1997]) one obtains a system on the Lie algebra $\mathfrak{g} = TG/G$, and the reduced Hamilton equations are the Hamilton equations obtained by means of the Lie–Poisson structure on \mathfrak{g}^* . The best known example of this type is probably the reduction for the rigid body in analytical mechanics: from the cotangent bundle $T^*\mathrm{SO}(3)$ of the group $\mathrm{SO}(3)$ playing the role of the configuration space to the linear Poisson structure on $\mathfrak{so}(3)^*$ —the dual of the Lie algebra $\mathfrak{so}(3)$. Similar situation appears for homogeneous spaces [Jurdjevic, 1997] and general principal bundles [Martínez, 2004, Martínez, 2007]. The reduced system lives on the bundle TP/G which is canonically a Lie algebroid, called the *Atiyah algebroid of P* , and the reduced Hamilton equations are associated with the linear Poisson structure on T^*P/G (equivalent to the presence of a Lie algebroid structure on TP/G). In this approach one obtains a version of the PMP, yet the Hamiltonian reduction seems to be purely computational and a big part of the geometry of the problem remains hidden.

The second approach, called the *Lagrangian reduction*, was introduced by Marsden and his collaborators (see for example [Cendra *et al.*, 1998]) in the context of analytical mechanics. Here, one uses the reduced data $f : P/G \times U \rightarrow TP/G$ and $l : P/G \times U \rightarrow \mathbb{R}$, and the reduced variations (homotopies) to obtain a reduced version of the Euler-Lagrange equations. In this approach it becomes clear that the reduction of the variational principle is not only the reduction of the data and geometrical structure, but also a reduction of variations (homotopies)—this is most clearly stated in [Cendra *et al.*, 1998] for the case of an invariant system on a Lie group. By means of the Lagrangian reduction one can obtain various results such as Euler-Poincaré equations and Hammel equations. Despite of this advantages, the Lagrangian reduction seems to be useful rather in mechanics than in control theory, as one requires the geometry of the set of controls U and controls itself being very regular (U should be at least an affine subspace of \mathbb{R}^n , and controls differentiable), so accepts no discontinuity, switch-on-switch-off controls, etc.

The main result

The aim of our work is to extend the fundamental theorem of optimal control—the PMP—to the setting of almost Lie (AL) algebroids—geometrical objects generalising Lie algebroids.

Since Lie algebroids are infinitesimal (reduced) objects of (local) Lie groupoids (like Lie algebras are for Lie groups), we are motivated mostly by the Lie groupoid \mathcal{G} —Lie algebroid $A(\mathcal{G})$ reduction. Obviously, a reduction of an invariant control system on a Lie groupoid should lead to a system on the associated Lie algebroid. An example of such a situation was discussed in the previous subsection, where an invariant control system on a principal bundle lead to a system on the associated Atiyah algebroid.

What is more, similarly to the scheme of the Lagrangian reduction, we should also reduce the variations (homotopies) from \mathcal{G} to $A(\mathcal{G})$. This will motivate the abstract definition of the homotopy of admissible paths on an AL algebroid (algebroid homotopy). Finally, reducing an invariant OCP on the Lie groupoid \mathcal{G} would not be complete without reducing the boundary conditions as well. The idea is to substitute fixed-end-points boundary conditions on \mathcal{G} by fixed-homotopy-class conditions. These two are closely related (see Chapter 4 for a detailed discussion) and equivalent if \mathcal{G} is α -simply connected (the homotopy class of a curve on a simply connected manifold is uniquely determined by its end-points). Now, since homotopies in \mathcal{G} correspond to algebroid homotopies in $\mathcal{A}(\mathcal{G})$, we can express the reduced boundary conditions in $A(\mathcal{G})$ as fixing the algebroid homotopy class of the trajectory of the reduced control system. A similar construction can be made also for more general boundary conditions.

At the end, for a general AL algebroid E , we can formulate an analog of the OCP which, in the case of an integrable algebroid $E = A(\mathcal{G})$, turns out to be an invariant OCP reduced from \mathcal{G} . Let us note that our understanding of algebroid homotopies and homotopy classes is closely related to that of Crainic and Fernandes [Crainic & Fernandes, 2003], where similar techniques were used to generalise the Third Theorem of Lie and integrate Lie algebroids. However, our framework is much more general, as we no longer remain in the smooth category.

To explain briefly the result, let us note that an AL algebroid is a vector bundle $\tau : E \rightarrow M$ together with a vector bundle map $\rho : E \rightarrow TM$ (*anchor*) and a skew-symmetric bilinear bracket $[\cdot, \cdot]$ on the space of sections of E which satisfy certain compatibility conditions. The algebroid structure on E is equivalent to the presence of a certain linear bi-vector field Π on the dual bundle E^* . Note that the bivector field Π defines the *Hamiltonian vector field* \mathcal{X}_H associated with any C^1 -function H on E^* , defined in the standard way as the contraction $\mathcal{X}_H = \iota_{dH}\Pi$.

Standard examples of AL (in fact Lie) algebroids are: the tangent bundle $E = TM \rightarrow M$ with $\rho = \text{id}_{TM}$ and the Lie bracket of vector fields, and a finite-dimensional real Lie algebra $E = \mathfrak{g}$ with the trivial anchor map (M is a single point in this case) and the Lie bracket on \mathfrak{g} . In the first case, Π is the canonical Poisson tensor on T^*M , whereas in the second—the Lie–Poisson structure on \mathfrak{g}^* . An example of an AL algebroid which is not a Lie algebroid is given by any real vector bundle with a smooth family of skew-symmetric bilinear (but not Lie) operations on its fibers.

On the bundle E we can consider *admissible paths*; i.e., bounded measurable maps $a : [t_0, t_1] \rightarrow E$ such that the projection $x(t) = \tau(a(t))$ of $a(t)$ onto M is absolutely continuous (AC) and $\dot{x}(t) = \rho(a(t))$ a.e. On admissible paths we have an equivalence relation $a \sim b$ interpreted as a reduction of homotopy equivalence (with fixed end-points). Note that equivalent paths need not to be defined on the same time interval. For an admissible path σ , we denote with $[\sigma]$ the equivalence class of σ .

A *control system* is defined by a continuous map $f : M \times U \rightarrow E$, where U is a topological space of control parameters such that, for each $u \in U$, the function $f(\cdot, u)$ is a section of class C^1 of the bundle E . Every *admissible control*, i.e., a bounded measurable path $u(t)$ in U , gives rise to an absolutely continuous path in M defined by the differential

equation

$$\dot{x}(t) = \rho(f(x(t), u(t))),$$

and to an admissible path $a(t) = f(x(t), u(t))$ covering $x(t)$. We will call $a(t)$ the *trajectory* of the control system and the pair $(x(t), u(t))$ —the *controlled pair*. An *optimal control problem* for this control system is associated with a fixed homotopy class $[\sigma]$ of an admissible path σ and a *cost function* $L : M \times U \rightarrow \mathbb{R}$. The problem is to find a controlled pair $(x(t), u(t))$ with $t \in [t_0, t_1]$ (the time interval is to be found as well) such that

$$\begin{aligned} \text{the integral } \int_{t_0}^{t_1} L(x(t), u(t)) dt \text{ is minimal among all controlled pairs } (x, u) \text{ for} \\ \text{which the } E\text{-homotopy class of the trajectory } f(x(t), u(t)) \text{ equals } [\sigma]. \end{aligned} \quad (\text{P})$$

Our main result is the following.

Theorem 1.1. *Let $(x(t), u(t))$, with $t \in [t_0, t_1]$, be a controlled pair solving the optimal control problem (P). Then there exists a curve $\xi : [t_0, t_1] \rightarrow E^*$ covering $x(t)$ and a constant $\xi_0 \leq 0$ such that*

- *the curve $\xi(t)$ is a trajectory of the time-dependent family of Hamiltonian vector fields \mathcal{X}_{H_t} , $H_t(x, \xi) := H(x, \xi, u(t))$, where*

$$H(x, \xi, u) = \langle f(x, u), \xi \rangle + \xi_0 L(x, u);$$

- *the control u satisfies the “maximum principle”*

$$H(x(t), \xi(t), u(t)) = \sup_{v \in U} H(x(t), \xi(t), v)$$

and $H(x(t), \xi(t), u(t)) = 0$ at every regular point t of u ;

- *if $\xi_0 = 0$, then the covector $\xi(t)$ is nowhere-vanishing.*

We have also developed a version of this result for general boundary conditions. These can be expressed by means of two smooth algebroid morphisms $\Phi_0 : TS_0 \rightarrow E$ and $\Phi_1 : TS_1 \rightarrow E$. In the integrable case $E = \mathcal{A}(\mathcal{G})$ it is convenient to think of Φ_0 and Φ_1 as of two smooth maps $\tilde{\Phi}_0 : S_0 \rightarrow \mathcal{G}$ and $\tilde{\Phi}_1 : S_1 \rightarrow \mathcal{G}$ reduced to $\mathcal{A}(\mathcal{G})$. Now we can formulate the *relative OCP* by substituting in the problem (P) the algebroid homotopy class $[\sigma]$ by the relative algebroid homotopy class $[\sigma] \bmod(\Phi_0, \Phi_1)$. Here the relative class can be understood as a reduction of a homotopy in \mathcal{G} , with end-points in the images $\text{Im } \tilde{\Phi}_0$ and $\text{Im } \tilde{\Phi}_1$, to the algebroid $\mathcal{A}(\mathcal{G})$.

For a solution of the problem described above we can repeat Theorem 1.1 with additional transversality conditions, namely, that the covectors $\xi(t_0)$ and $\xi(t_1)$ annihilate the images $\text{Im } \Phi_0$ and $\text{Im } \Phi_1$, respectively.

Discussion of the main result

The above result looks quite similar to the standard PMP. Indeed, in the case $E = TM$ we obtain the PMP. The only difference is that the fixed-end-point boundary condition are substituted by the fixed-homotopy-class condition. However, this makes no essential difference, as is discussed in detail in Chapter 4. For the case of an integrable Lie algebroid $E = A(\mathcal{G})$ our version of the PMP can be understood as a general reduction scheme for invariant OCPs on Lie groupoids. In particular, the theorem covers the known results on Hamiltonian reduction of Jurdjevic [Jurdjevic, 1997] and Martínez [Martínez, 2004, Martínez, 2007], and Lagrangian reduction [Cendra *et al.*, 1998] (see Chapter 6 for details). It is, however, worth mentioning that in our approach the reduced and the unreduced PMPs are parts of the same universal formalism. Roughly speaking, we have generalised the geometrical context in which the PMP can be used. The technical setting remains quite general—we work with bounded measurable controls and AC base trajectories. Moreover, since AL algebroids do not come, in general, from reductions, our result admits a wider spectrum of possible applications. An attempt in this direction can be found in the last example of Chapter 6. Finally, note that a version of Theorem 1.1 for general boundary conditions admits arbitrary algebroid morphisms $\Phi_0 : TS_0 \rightarrow E$ and $\Phi_1 : TS_1 \rightarrow E$. In the integrable case $E = A(\mathcal{G})$ these correspond to arbitrary smooth maps $\tilde{\Phi}_0 : S_0 \rightarrow \mathcal{G}$ and $\tilde{\Phi}_1 : S_1 \rightarrow \mathcal{G}$. On the other hand, in literature, when speaking about general boundary conditions one usually restricts attention to immersions only.

The original contributions of the author includes:

- a detailed study of the notion of an algebroid homotopy and algebroid homotopy classes in Chapter 3:
 - The definition of the Lie algebroid homotopy appeared in [Crainic & Fernandes, 2003]. It was given in terms of time-dependent algebroid sections and connections and though was not very intuitive. We extended the notion of the algebroid homotopy to almost Lie algebroids, extended it to measurable class, and reformulated the definition to emphasise the similarities with the standard notion of homotopy,
 - We introduced the notion of a relative algebroid homotopy class.
 - We gave a new interpretation of algebroid homotopies in terms of a Stokes-like formula and also extended the well-known interpretation of algebroid homotopies as reduced homotopies of a groupoid to the measurable class (Theorem 3.6).
 - We ask a question about existence and uniqueness of algebroid homotopies. A uniqueness result (Lemma 3.3) is a simple consequence of certain results from the theory of differential equations. On the other hand, the existence is strongly connected with the axioms of AL algebroid. We prove Lemma 3.12 which states that only for AL algebroids every sufficiently regular one-parameter

family of admissible paths generates an algebroid homotopy (for a given initial-point algebroid homotopy). This result, which in its infinitesimal and smooth version appeared earlier in [Grabowska & Grabowski, 2008], distinguishes AL algebroids from more general objects of similar nature (skew-algebroids or general algebroids).

- We study further properties of algebroid homotopies in Section 3.2. These include Lemma 3.17, which shows the behaviour of algebroid homotopy classes under reparametrisation, and Lemma 3.15, which compares algebroid homotopies with and without fixed end-points.
- formulating the OCPs in the language of AL algebroids in Chapter 4:
 - We proposed to express boundary conditions of the OCPs in terms of algebroid homotopies.
 - We gave an interpretation of these new OCPs and, in particular, studied in detail their relation to standard OCPs.
 - We proposed to express general boundary conditions in the OCP in terms of algebroid morphisms rather than submanifolds.
- formulating a version of the PMP in the language of AL algebroids for fixed end-points and general boundary conditions (Theorems 5.1 and 5.2);
- proving these theorems in Chapters 7–9:
 - The proof, in principle, imitates the argument of Boltyanskii [Pontryagin *et al.*, 1962]. There are, however, technical difficulties connected with using the language of AL algebroids. These appeared mostly in two places. In the proof of Theorem 7.4 we used reparametrisation and composition of algebroid homotopies to study the impact of needle variations on the trajectories of a control system. In Lemma 8.1 to prove the existence of an admissible path realising a certain algebroid homotopy class we had to pass through infinite-dimensional Banach spaces. The reason for that is the following: AL algebroids are, in general, not integrable, and hence homotopy classes cannot be represented by points on a finite dimensional manifold (they are just cosets in a big space of curves). Our idea was to semi-parametrise these classes by a finite-dimensional space and reduce the reasoning to a finite dimensional topological problem.
 - Moreover, the standard proof [Pontryagin *et al.*, 1962] like most of the other proofs in literature (perhaps apart from [Barbero-Liñán & Muñoz-Lecanda, 2009]) contains smaller or greater gaps. We put much effort to explain all the details and make the reasoning self-contained.

A significant part of this work is based on [Grabowski & Józwickowski, 2011]. In this article we concentrated only on OCPs with fixed-end-point boundary conditions. Therefore all parts concerning general boundary conditions, in particular Definition 3.10, Theorem 5.2 and its proof including Lemma 8.2 and Section 9.2, and parts concerning an interpretation

of relative algebroid homotopies and an interpretation of the OCP (P rel) (in Chapter 4), has not been published before. Also broad parts of [Grabowski & Józwiński, 2011] were reformulated to make the argument more understandable.

Organisation of the manuscript

The first major part of this work is intended to give all important definitions and motivations which allow to define OCPs on an AL algebroid in Chapter 4 and finally state our main results in Chapter 5. We start with a brief introduction of AL algebroids in Chapter 2. In Chapter 3 we concentrate on algebroid homotopies which are crucial in our work. Much effort was made to give a satisfactory definition in both smooth and measurable setting, and later to motivate this definition, mainly by Lie groupoid—Lie algebroid reduction arguments. We also derive all properties of algebroid homotopies which will be used later in the proof of our main results. Finally, in Chapter 4, we define and motivate algebroid OCPs. Much attention is paid to algebroid homotopies naturally associated with a control system on an AL algebroid. This leads to the notion of a parallel transport.

In Chapter 6 we derive some known results on reduction in optimal control theory and variational calculus by means of our result. In particular we formulate the version of the PMP for invariant OCPs on principal bundles and use it to study the example of the falling cat problem of [Montgomery, 1990]. Some attention is paid to the problems of the calculus of variations on principal bundles. We obtain the results on Lagrangian reduction, Hammel equations and Euler–Poincaré equations as a special case. We also derive the generalised Euler-Lagrange equations on a general AL algebroid.

The second major part, consisting of Chapters 7–9, contains the proof of Theorem 1.1. In Chapter 7 we define needle variations and a cone \mathbf{K}_τ^u of infinitesimal variations of the trajectory of the control system. The geometry of this cone is studied in detail in Chapter 9, using technical results proved in Chapter 8. Then we can follow [Pontryagin *et al.*, 1962] to derive the necessary conditions for optimality from the geometric properties of \mathbf{K}_τ^u along the optimal trajectory.

Parallel to the proof of Theorem 1.1 we prove its version with general boundary conditions. Usually this requires just a minor modification of the arguments used. We decided to give two proofs in spite of the fact that Theorem 1.1 is just a special case of the version with general boundary conditions. We believe that in this way the already complicated reasoning is easier to follow. Moreover, this is the typical way the proof of the PMP is presented in literature.

In the main part of this work we assume that the reader is familiar with basics of control theory, geometry of convex sets, topology, theory of ODEs in the sense of Carathéodory, and basics of the theory of Lie groupoids. However, the reader who is not confident with these topics can find necessary information in Appendixes A–D (we give the references when necessary). We believe that our presentation is self-contained. The Appendixes contain also some minor technical results which are used in the argument, yet their derivation in the main text would make the presentation less clear.

Almost Lie algebroids

This chapter is concerned with some basic definitions and constructions from the theory of algebroids. We begin with the definition of a skew-algebroid and an almost Lie algebroid as a special case. Later we introduce the notion of a Hamiltonian vector field and the complete lift of an algebroid section. The characterisation of skew-algebroids in terms of exterior differential operators is used to define a morphism of algebroids. This, in turn, leads to the notion of an admissible path. We end this chapter with the construction of the product of two algebroids.

Let us note that many aspects of the theory of algebroids are not present in this introductory chapter. The interested reader should confront [Mackenzie, 2005, Mackenzie, 1987, Grabowski & Urbański, 1999, Silva & Weinstein, 1999].

Differentiable manifolds and vector bundles

In this work we use the following notation and conventions of differential geometry. By M we denote a smooth n -dimensional manifold, by $\tau_M : TM \rightarrow M$ the tangent vector bundle, and by $\pi_M : T^*M \rightarrow M$ the cotangent vector bundle of M . When passing to a local description we will use a coordinate system (x^a) , $a = 1, \dots, n$ in M . We have the induced (adapted) coordinate systems (x^a, \dot{x}^b) in TM and (x^a, p_b) in T^*M .

More generally, let $\tau : E \rightarrow M$ be a vector bundle, and let $\pi : E^* \rightarrow M$ be the dual bundle. Choose (e_1, \dots, e_m) — a basis of local sections of $\tau : E \rightarrow M$, and let (e_*^1, \dots, e_*^m) be the dual basis of local sections of $\pi : E^* \rightarrow M$. We have the induced coordinate systems: (x^a, y^i) , $y^i = \iota(e_*^i)$ in E , and (x^a, ξ_i) , $\xi_i = \iota(e_i)$ in E^* , where the linear functions $\iota(e)$ are given by the canonical pairing $\iota(e)(v_x) = \langle e(x), v_x \rangle$. The null section of $\tau : E \rightarrow M$ will be denoted by θ , and θ_x will stand for the null vector at point $x \in M$.

In this work the summation convention is assumed.

Almost Lie algebroids

Definition 2.1. Let M be a manifold and $\tau : E \rightarrow M$ a vector bundle over M . A *skew-algebroid structure* on E is a vector bundle morphism $\rho : E \rightarrow TM$ over M , called the *anchor map*, and a skew-symmetric bilinear bracket $[\cdot, \cdot] : \text{Sec}(E) \times_M \text{Sec}(E) \rightarrow \text{Sec}(E)$ on (local) sections of τ , which satisfies the Leibniz rule

$$[X, f \cdot Y] = f[X, Y] + \rho(X)(f)Y \quad (2.1)$$

for every $X, Y \in \text{Sec}(E)$ and $f \in C^\infty(M)$.

If, additionally, the anchor map is an algebroid morphism, i.e.,

$$\rho([X, Y]) = [\rho(X), \rho(Y)]_{TM}, \quad (2.2)$$

we will speak of an *almost Lie algebroid* (AL algebroid briefly).

If, in addition to (2.1) and (2.2), the bracket satisfies the Jacobi identity (in other words, the pair $(\text{Sec}(E), [\cdot, \cdot])$ is a Lie algebra), we speak of a *Lie algebroid*.

In local coordinates (x^a, y^i) , introduced at the beginning of this chapter the structure of an algebroid on E can be described in terms of local function $\rho_i^a(x)$ and $c_{jk}^i(x)$ on M given by

$$\rho(e_i) = \rho_i^a(x) \partial_{x^a} \quad \text{and} \quad [e_i, e_j] = c_{ij}^k(x) e_k.$$

The skew-symmetry of the algebroid bracket results in the skew-symmetry of c_{jk}^i in lower indices, whereas condition (2.2) reads as

$$\left(\frac{\partial}{\partial x^b} \rho_k^a(x) \right) \rho_j^b(x) - \left(\frac{\partial}{\partial x^b} \rho_j^a(x) \right) \rho_k^b(x) = \rho_i^a(x) c_{jk}^i(x).$$

In the context of mechanics it is convenient to think about an algebroid as a generalisation of the tangent bundle. An element $a \in E$ has the interpretation of a generalized velocity with actual velocity $v \in TM$ obtained by applying the anchor map $v = \rho(a)$. The kernel of the anchor map represents inner degrees of freedom.

A basic example of a skew-algebroid structure is the tangent bundle TM of a manifold M with the standard Lie bracket and $\rho = \text{id}_{TM}$. We will refer to this structure as to a *tangent algebroid*. Another natural example is a finite-dimensional real Lie algebra \mathfrak{g} considered as a vector bundle over a single point with its Lie bracket and the trivial anchor.

Natural examples of skew-algebroids are associated with systems with symmetries. For instance, the Lie algebra \mathfrak{g} of a Lie group G can be understood as a reduction of the tangent bundle TG by the left (or right) action of G . Similarly, for a principal bundle $G \rightarrow P \rightarrow M$, the reduced bundle $TP/G \rightarrow M$ has the structure of an *Atiyah algebroid*. The Atiyah algebroid is a common generalisation of TM and \mathfrak{g} . This example is discussed in more details in Appendix A.2. More generally, every Lie groupoid \mathcal{G} has an associated Lie algebroid $\mathcal{A}(\mathcal{G})$ which can be interpreted as a reduction of a subbundle of the tangent algebroid $T\mathcal{G}$ by the right (or left) action of \mathcal{G} . This example is discussed in Appendix A.1.

All the above are examples of Lie algebroids. Natural examples of skew-algebroids which are not Lie can be associated with nonholonomically constrained mechanical systems [Grabowski *et al.*, 2009].

Hamiltonian vector fields and tangent lifts

Let us now describe some geometric constructions associated with the structure of a skew-algebroid .

It can be shown (cf. [Grabowski & Urbański, 1999, Grabowski & Urbański, 1997]) that the presence of the structure of a skew-algebroid on E is equivalent to the existence of a linear bivector field Π_{E^*} on E^* . In local coordinates, (x^a, ξ_i) on E^* , it is given by

$$\Pi_{E^*} = c_{ij}^k(x) \xi_k \partial_{\xi_i} \wedge \partial_{\xi_j} + \rho_i^b(x) \partial_{\xi_i} \wedge \partial_{x^b}. \quad (2.3)$$

The linearity of Π_{E^*} means that the corresponding mapping $\tilde{\Pi} : T^*E^* \rightarrow TE^*$ is a morphism of double vector bundles (cf. [Konieczna & Urbański, 1999, Grabowski & Rotkiewicz, 2009]). The tensor Π_{E^*} is well recognised in the standard situations: for the tangent algebroid structure on TM , it is the canonical Poisson structure on T^*M dual to the canonical symplectic structure, whereas for a Lie algebra \mathfrak{g} , it is the Lie–Poisson structure $\Pi_{\mathfrak{g}^*}$ on \mathfrak{g}^* . Actually, E is a Lie algebroid if and only if Π_{E^*} is a Poisson tensor.

Now we can introduce the notion of a Hamiltonian vector field on E^* . Let, namely, $h : E^* \rightarrow \mathbb{R}$ be any C^1 -function. We define the *Hamiltonian vector field* \mathcal{X}_h in an obvious way: $\mathcal{X}_h = \iota_{dh} \Pi_{E^*}$. In local coordinates,

$$\mathcal{X}_h(x, \xi) = \rho_i^a(x) \frac{\partial h}{\partial \xi_i}(x, \xi) \partial_{x^a} + \left(c_{ji}^k(x) \xi_k \frac{\partial h}{\partial \xi_j}(x, \xi) - \rho_i^a(x) \frac{\partial h}{\partial x^a}(x, \xi) \right) \partial_{\xi_i}. \quad (2.4)$$

Another geometrical construction in the skew-algebroid setting is the *complete lift of an algebroid section* (cf. [Grabowski & Urbański, 1999, Grabowski & Urbański, 1997]). For every C^1 -section $X = f^i(x) e_i \in \text{Sec}(E)$ we can construct canonically a vector field $d_T(X) \in \text{Sec}(TE)$ which, in local coordinates, reads as

$$d_T(X)(x, y) = f^i(x) \rho_i^a(x) \partial_{x^a} + \left(y^i \rho_i^a(x) \frac{\partial f^k}{\partial x^a}(x) + c_{ij}^k(x) y^i f^j(x) \right) \partial_{y^k}. \quad (2.5)$$

The vector field $d_T(X)$ is linear w.r.t. the vector bundle structure $T\tau : TE \rightarrow TM$ (the above equation is linear w.r.t. y^i).

Consider the Hamiltonian vector field $\mathcal{X}_{\iota(X)}$ associated with a linear function $\iota(X)(\cdot) = \langle X, \cdot \rangle_\tau$ on E^* . It turns out that fields $\mathcal{X}_{\iota(X)}$ and $d_T(X)$ are related by

$$\langle d_T(X), \mathcal{X}_{\iota(X)} \rangle_{T\tau} = 0, \quad (2.6)$$

where $\langle \cdot, \cdot \rangle_{T\tau} : TE \times_{TM} TE^* \rightarrow \mathbb{R}$ is the canonical pairing, being the tangent map of $\langle \cdot, \cdot \rangle_\tau : E \times_M E^* \rightarrow \mathbb{R}$ (in local coordinates, $\langle (x, y, \dot{x}, \dot{y}), (x, \xi, \dot{x}, \dot{\xi}) \rangle_{T\tau} = \dot{y}^j \xi_j + y^j \dot{\xi}_j$).

Cartan Calculus

The existence of a skew-algebroid structure on E is equivalent to the existence of exterior differential (de Rham) operators

$$d_E : \text{Sec}(\Lambda^k E^*) \rightarrow \text{Sec}(\Lambda^{k+1} E^*), \quad k = 0, 1, \dots,$$

defined by a straightforward generalisation of the Cartan formula

$$\begin{aligned} d_E \omega(a_0, a_1, \dots, a_k) &= \sum_{i=0}^k (-1)^i \rho(a_i) \omega(a_0, \dots, \check{a}_i, \dots, a_k) \\ &+ \sum_{1 \leq i < j \leq k} (-1)^{i+j} \omega([a_i, a_j], a_0, \dots, \check{a}_i, \dots, \check{a}_j, \dots, a_k), \end{aligned}$$

for $\omega \in \text{Sec}(\Lambda^k E^*)$ and $a_0, a_1, \dots, a_k \in \text{Sec}(E)$.

These operators, in general, needs not be cohomological. In fact, E is a Lie algebroid if and only if $d_E^2 = 0$. AL algebroids, in turn, can be characterized by the condition that $d_E^2 f = 0$ for every $f \in C^\infty(M) = \text{Sec}(\Lambda^0 E^*)$.

Morphisms

The above concept of the de Rham derivative allows one to give a simple definition of a morphism of skew-algebroids. Namely, given skew-algebroids $\tilde{\tau} : \tilde{E} \longrightarrow \tilde{M}$ and $\tau : E \longrightarrow M$, a bundle map $\Phi : \tilde{E} \longrightarrow E$ over $\varphi : \tilde{M} \longrightarrow M$ is a *skew-algebroid morphism* if it is compatible with the exterior derivative:

$$\Phi^* d_E \theta = d_{\tilde{E}} \Phi^* \theta, \quad \text{for every } \theta \in \text{Sec}(\Lambda^k E^*). \quad (2.7)$$

Note that a vector bundle map Φ does not, in general, induce any map on sections of \tilde{E} , while the pull-back Φ^* of sections of E^* is always well defined.

Introduce local coordinates $(\tilde{x}^\alpha, \tilde{y}^\iota)$ and (x^a, y^i) and structure functions $\tilde{\rho}_\iota^\alpha(\tilde{x})$, $\tilde{c}_{\kappa\mu}^\iota(\tilde{x})$ and $\rho_i^a(x)$, $c_{km}^i(x)$ on \tilde{E} and E , respectively. The condition that $\Phi \sim (\Phi_\iota^i, \varphi^a)$ is an algebroid morphism reads as

$$\begin{aligned} \Phi_\kappa^i(\tilde{x}) \rho_i^a(\varphi(x)) &= \tilde{\rho}_\kappa^\alpha(\tilde{x}) \frac{\partial \varphi^a(\tilde{x})}{\partial \tilde{x}^\alpha}, \\ \tilde{\rho}_\kappa^\alpha(\tilde{x}) \frac{\partial \Phi_\lambda^i(\tilde{x})}{\partial \tilde{x}^\alpha} - \tilde{\rho}_\lambda^\alpha(\tilde{x}) \frac{\partial \Phi_\kappa^i(\tilde{x})}{\partial \tilde{x}^\alpha} &= c_{jk}^i(\varphi(\tilde{x})) \Phi_\kappa^j(\tilde{x}) \Phi_\lambda^k(\tilde{x}). \end{aligned} \quad (2.8)$$

Admissible paths

Consider an algebroid morphism $\text{T}\mathbb{R}|_I \longrightarrow E$, where $I = [t_0, t_1] \subset \mathbb{R}$ is an interval. Every such map is uniquely determined by the image of the canonical section (t, ∂_t) of $\text{T}\mathbb{R}$ being a smooth curve $a(t)$ in E over the base path $x(t)$ in M . Condition (2.7) reads as

$$\rho(a(t)) = \dot{x}(t) \quad \text{for every } t \in I. \quad (2.9)$$

This means that the anchor map coincides with the tangent prolongation of the projection $x(t) = \tau(a(t))$. The curves which satisfy (2.9) will be called *admissible*. In fact, (2.9) also makes sense for non-smooth maps. From now on, by an *admissible path on E* (or briefly *E -path*) we shall mean a bounded measurable map $a : I \longrightarrow E$ over an absolutely continuous (AC) base path $x = \tau \circ a : I \longrightarrow M$ such that (2.9) is satisfied a.e. in I . In such a case we will speak of *measurable E -paths*. For more information on measurable functions see Section B.1. Observe that from (2.8) it follows that a morphism of algebroids maps admissible paths into admissible paths.

To explain the meaning of admissible curves, observe that in the case of the tangent algebroid $\text{T}M$ admissible curves are precisely the tangent lifts of base curves. We will show later (cf. Theorem 3.6) that if an algebroid E is integrable, admissible curves come from a reduction of real curves in a Lie groupoid integrating E .

Finally, we can introduce the concept of *composition of measurable E -paths*. Let $a : [t_0, t_1] \rightarrow E$ and $\bar{a} : [t_1, t_2] \rightarrow E$ be two measurable E -paths with base paths $x = \tau \circ a$ and $\bar{x} = \tau \circ \bar{a}$, respectively. Assume that $x(t_1) = \bar{x}(t_1)$ (such paths will be called *composable*). Clearly, the map $\tilde{a} : [t_0, t_2] \rightarrow E$ defined by

$$\tilde{a}(t) := \begin{cases} a(t) & \text{for } t \leq t_1, \\ \bar{a}(t) & \text{for } t > t_1 \end{cases}$$

is another measurable E -path covering the AC curve

$$\tilde{x}(t) := \begin{cases} x(t) & \text{for } t \leq t_1, \\ \bar{x}(t) & \text{for } t > t_1. \end{cases}$$

This new E -path will be called the *composition* of a and \bar{a} and will be denoted by $\tilde{a} = a \circ \bar{a}$.

The product of skew-algebroids

Given two skew-algebroids $(\tau_1 : E_1 \rightarrow M_1, \rho_1, [\cdot, \cdot]_1)$ and $(\tau_2 : E_2 \rightarrow M_2, \rho_2, [\cdot, \cdot]_2)$ we can define a skew-algebroid structure on the product bundle $\tau = \tau_1 \times \tau_2 : E_1 \times E_2 \rightarrow M_1 \times M_2$. The anchor will simply be $\rho = \rho_1 \times \rho_2 : E_1 \times E_2 \rightarrow TM_1 \times TM_2 \approx T(M_1 \times M_2)$. The bracket can be defined by equalities

$$\begin{aligned} [p_1^* X_1, p_1^* Y_1] &= p_1^* [X_1, Y_1]_1, \\ [p_2^* X_2, p_2^* Y_2] &= p_2^* [X_2, Y_2]_2, \\ [p_1^* X_1, p_2^* Y_2] &= \theta, \end{aligned}$$

where $X_1, Y_1 \in \text{Sec}(E_1)$ and $X_2, Y_2 \in \text{Sec}(E_2)$ are sections, $p_1 : E_1 \times E_2 \rightarrow E_1$ and $p_2 : E_1 \times E_2 \rightarrow E_2$ are canonical vector bundle projections, and θ is a null section of τ . The above equalities can be extended to arbitrary sections by linearity and the Leibniz rule (2.1). Clearly, the canonical projections $E_1 \times E_2 \rightarrow E_i$, with $i = 1, 2$, are algebroid morphisms, and if E_1 and E_2 are almost Lie, then so is their product.

The local coordinate description of the product $E_1 \times E_2$ is very simple. If (x^a, y^i) and $(\tilde{x}^\alpha, \tilde{y}^\iota)$ are local coordinates on E_1 and E_2 , respectively, we can introduce natural coordinates $(X^A, Y^I) = (x^a, \tilde{x}^\alpha, y^i, \tilde{y}^\iota)$ on $E_1 \times E_2$. The structure functions $C_{JK}^I(X)$ and $R_I^A(X)$ in these coordinates are trivial on mixed-type terms ($R_i^\alpha = C_{j\kappa}^\iota = 0$, etc.) and the same as the structure functions of E_1 and E_2 on simple-type terms ($C_{jk}^i(x, \tilde{x}) = c_{jk}^i(x)$, $R_\iota^\alpha(x, \tilde{x}) = \rho_\iota^\alpha(\tilde{x})$, etc.).

Homotopies of admissible paths

The notion of the homotopy of E -paths (E -homotopy) is crucial in this work. It will be our main tool to define the boundary conditions for optimal control problem on AL algebroids in Chapter 4. In this chapter we give a definition of E -homotopy and study its basic properties.

In the first section much attention is put on interpreting E -homotopy. In particular, if $E = \mathcal{A}(\mathcal{G})$ is a Lie algebroid of a Lie groupoid \mathcal{G} , we show an equivalence of E -homotopies and true homotopies in \mathcal{G} (Theorem 3.6). We also interpret E -homotopies by a Stokes-like formula. Finally, we introduce an important notion of E -homotopy classes and E -homotopy classes relative to a pair of algebroid morphisms.

In the second section we study properties of E -homotopies. We prove an important Lemma 3.12 which states that AL algebroids are characterised by the property that a one-parameter family of E -paths establishes an E -homotopy. This result explains why AL algebroids are objects of our primary interest rather than a smaller class of Lie algebroids or a more general class of skew-algebroids. Later we prove Lemma 3.15 which compares E -homotopies with and without fixed end-points. Finally, the behaviour of E -homotopy classes under reparametrisation (Lemma 3.17) is studied.

3.1 The E -homotopy and its interpretation

The definition of E -homotopy will be given in two steps. First, we will describe the smooth case and later generalise the concept to measurable E -paths, more suitable in control theory. The rest of this section is concerned with giving a convincing motivation and interpretation of the E -homotopy. We give an interpretation in terms of a Stokes-like formula, and later show that E -homotopies on an integrable algebroid $\mathcal{A}(\mathcal{G})$ correspond to true homotopies in the groupoid \mathcal{G} integrating $\mathcal{A}(\mathcal{G})$. The last interpretation will be crucial in Chapter 4 to motivate the definition of an optimal control problem on an AL algebroid.

Definition 3.1. Let $a_0, a_1 : I = [t_0, t_1] \rightarrow E$ be two smooth admissible paths. An *algebroid homotopy in E* (or *E -homotopy* briefly) between a_0 and a_1 is a pair of maps $a, b : [t_0, t_1] \times [0, 1] \rightarrow E$, over the same base map $x : [t_0, t_1] \times [0, 1] \rightarrow M$, with

$a_0(\cdot) = a(\cdot, 0)$ and $a_1(\cdot) = a(\cdot, 1)$, such that

$$t \mapsto a(t, s) \quad \text{is admissible for every } s \in [0, 1], \quad (3.1a)$$

$$s \mapsto b(t, s) \quad \text{is admissible for every } t \in [t_0, t_1], \quad (3.1b)$$

and, moreover, a and b satisfy a system of differential equations given in local coordinates (x^a, y^i) in E by

$$\partial_t b^i(t, s) - \partial_s a^i(t, s) = c_{jk}^i(x(t, s)) b^j(t, s) a^k(t, s). \quad (3.2)$$

The E -paths $b_0(s) := b(t_0, s)$ and $b_1(s) := b(t_1, s)$ will be called *initial-point* and *final-point* E -homotopies, respectively. We will say that E -homotopy (a, b) has *fixed end-points* if $b_0 \equiv \theta_{x(t_0)}$ and $b_1 \equiv \theta_{x(t_1)}$.

Having in mind that admissible paths are in a 1-1 correspondence with algebroid morphisms $A : \mathbb{T}\mathbb{R}|_I \longrightarrow E$ (an admissible path is the A -image of the canonical section (t, ∂_t) of $\mathbb{T}\mathbb{R}$ — compare Chapter 2), we may define E -homotopy in an equivalent way. An E -homotopy between two smooth algebroid morphisms $A_0, A_1 : \mathbb{T}\mathbb{R}|_I \longrightarrow E$ (corresponding to E -paths a_0 and a_1) is an algebroid morphism

$$H : \mathbb{T}\mathbb{R}|_I \times \mathbb{T}\mathbb{R}|_{[0,1]} \longrightarrow E,$$

such that $A_0(\cdot) = H(\cdot, \theta_0)$ and $A_1(\cdot) = H(\cdot, \theta_1)$, where $\theta_0 \in T_0\mathbb{R}$ and $\theta_1 \in T_1\mathbb{R}$ are null vectors. The equivalence with the previous definition can be seen as follows. The map H is determined by the images of two canonical sections (∂_t, θ_s) and (θ_t, ∂_s) . We can define $a(t, s) := H(\partial_t, \theta_s)$ and $b(t, s) := H(\theta_t, \partial_s)$. Now conditions (2.8) for H to be an algebroid morphism, translated to the language of a and b , are precisely (3.1a), (3.1b) and (3.2). Note that this alternative formulation agrees with the notion of the homotopy of Lie algebroid morphisms as introduced by [Kubarski, 1994].

The notion of an E -homotopy can be also extended to measurable setting.

Definition 3.2. Consider two bounded measurable admissible paths $a_0, a_1 : [t_0, t_1] \longrightarrow E$. An *algebroid homotopy in E* (or *E -homotopy* briefly) between a_0 and a_1 is a pair of bounded measurable maps (w.r.t. both variables separately) $a, b : [t_0, t_1] \times [0, 1] \longrightarrow E$, over the same ACB base map $x : [t_0, t_1] \times [0, 1] \longrightarrow M$, such that

- $a_0(\cdot) = a(\cdot, 0)$ and $a_1(\cdot) = a(\cdot, 1)$ are well-defined trace values,
- (3.1a) and (3.1b) hold in a measurable sense,
- the pair (a, b) is a weak solution of (3.2) with a well-defined trace (see Appendix B.3), that is

$$\begin{aligned} & \iint_{I \times [0,1]} \left[b^i(t, s) \partial_t \psi_i(t, s) - a^i(t, s) \partial_s \psi_i(t, s) + c_{jk}^i(x(t, s)) b^j(t, s) a^k(t, s) \psi_i(t, s) \right] dt ds \\ &= \int_{[0,1]} \left[b^i(0, s) \psi_i(0, s) - b^i(1, s) \psi_i(1, s) \right] ds - \int_I \left[a^i(t, 0) \psi_i(t, 0) - a^i(t, 1) \psi_i(t, 1) \right] dt \end{aligned} \quad (3.3)$$

holds for every family of functions $\psi_i \in C^\infty(I \times [0, 1]; \mathbb{R})$.

Note that considering only W -solutions of (3.2) would not be enough, since otherwise the boundary paths $a_0(t)$, $a_1(t)$, $b_0(s)$, and $b_1(s)$ would not be well defined. For more information on W - and WT -solutions see Appendix B.3. The notion of the *initial-point* and the *final-point E -homotopy*, as well as the *E -homotopy with fixed end-points*, also remains valid in this new setting.

From now on, by an E -homotopy we will mean a pair of maps (a, b) in the sense of Definition 3.2.

Observe that any two measurable maps $a, b : I \times [0, 1] \rightarrow E$ over the same AC base map $x : I \times [0, 1] \rightarrow M$ define a measurable bundle map $H : \mathbb{T}\mathbb{R}|_I \times \mathbb{T}\mathbb{R}|_{[0,1]} \rightarrow E$ (that is, a measurable map linear on fibers), where $H(\partial_t, \theta_s) = a(t, s)$ and $H(\theta_t, \partial_s) = b(t, s)$.

Like in the case of E -paths, also E -homotopies allow a natural notion of *composition*. Let, namely, $a, b : I \times [0, 1] \rightarrow E$ over x , and $\bar{a}, \bar{b} : J \times [0, 1] \rightarrow E$ over \bar{x} (where $I = [t_0, t_1]$ and $J = [t_1, t_2]$) be two E -homotopies. Assume that the final-point E -homotopy of the first and the initial-point E -homotopy of the second coincide; i.e., $b(t_1, s) = \bar{a}(t_1, s)$ a.e. (hence $x(t_1, s) = \bar{x}(t_1, s)$, so $a(\cdot, s)$ and $\bar{a}(\cdot, s)$ are composable for every $s \in [0, 1]$). The maps $\tilde{a}, \tilde{b} : I \cup J \times [0, 1] \rightarrow E$ defined as

$$\tilde{a}(t, s) = \begin{cases} a(t, s) & \text{for } t \leq t_1, \\ \bar{a}(t, s) & \text{for } t > t_1 \end{cases}$$

and

$$\tilde{b}(t, s) = \begin{cases} b(t, s) & \text{for } t \leq t_1, \\ \bar{b}(t, s) & \text{for } t > t_1 \end{cases}$$

clearly form an E -homotopy joining $a(\cdot, 0) \circ \bar{a}(\cdot, 0)$ and $a(\cdot, 1) \circ \bar{a}(\cdot, 1)$. The initial-point E -homotopy is $b(t_0, \cdot)$, while the final-point E -homotopy is $\bar{b}(t_2, \cdot)$.

Uniqueness of E -homotopies

As a direct consequence of the definition of an E -homotopy and Lemma B.19 we get the following result.

Lemma 3.3 (uniqueness of E -homotopies). *Let $a : I \times [0, 1] \rightarrow E$ be a bounded measurable map covering $x : I \times [0, 1] \rightarrow M$ such that $t \mapsto a(t, s)$ is admissible for every s . Then there exists at most one bounded measurable map $b : I \times [0, 1] \rightarrow E$ covering x such that (a, b) is an E -homotopy with a given initial-point E -homotopy $b(t_0, s) = b_0(s)$.*

The E -homotopy via Stokes theorem

We shall now give another, more geometrical, description of an E -homotopy by means of a Stokes-like formula. First, we will introduce the notion of an integral of an E - k -form, i.e., an element $\omega \in \text{Sec}(\Lambda^k E^*)$, over a bundle morphism $\Phi : \mathbb{T}N \rightarrow E$. We define

$$\int_{\Phi(N)} \omega := \int_N \Phi^* \omega,$$

where the last integral is the standard integration of the differential k -form $\Phi^*\omega$ on the manifold N . Now, if N is a manifold with boundary ∂N , we define

$$\int_{\partial\Phi(N)} \omega := \int_{\partial N} \Phi^*\omega.$$

Observe that in case $\Phi : TN \longrightarrow TM$ is the tangent lift of a diffeomorphism $\varphi : N \longrightarrow M$, the above definitions coincide with the standard concept of differential form integration. The morphism Φ need not be differentiable. Since, given local coordinates $N \supset V \xrightarrow{\psi} V' \subset \mathbb{R}^n \ni (y^1, \dots, y^n)$ on N ,

$$\int_V \Phi^*\omega = \int_{V' \subset \mathbb{R}^n} \omega(\Phi(\partial_{y^1}), \dots, \Phi(\partial_{y^n})) dy^1 \cdots dy^n,$$

we shall require only that Φ maps smooth sections of TN into bounded measurable sections of E .

Now assume that $\Phi : T\mathbb{R}|_I \times T\mathbb{R}|_{[0,1]} \longrightarrow E$ over $\varphi : \mathbb{R}|_I \times \mathbb{R}|_{[0,1]} \longrightarrow M$ is a bundle map defined by means of measurable maps $a(t, s)$ and $b(t, s)$ as in the definition of an E -homotopy. Assume, moreover, that conditions (3.1a) and (3.1b) are satisfied. Take any E -1-form $\alpha \in \text{Sec}(E^*)$; in local coordinates, $\alpha \sim (x^a, \alpha_i(x))$. Now

$$\begin{aligned} \int_{\Phi} d_E \alpha &= \iint_{I \times [0,1]} d_E \alpha(\Phi(\partial_t), \Phi(\partial_s)) dt ds = \iint_{I \times [0,1]} d_E \alpha(a(t, s), b(t, s)) dt ds = \\ &= \iint_{I \times [0,1]} \left(\rho_i^a(x) a^i \frac{\partial \alpha_j}{\partial x^a} b_j - \rho_i^a(x) b^i \frac{\partial \alpha_j}{\partial x^a} a_j - \alpha_i c_{jk}^i(x) a^j b^k \right) dt ds. \end{aligned}$$

Having in mind that $\rho_i^a(x(t, s)) a^i = \partial_t x^a(t, s)$ and $\rho_i^a(x(t, s)) b^i = \partial_s x^a(t, s)$, and defining $\tilde{\alpha}_i(t, s) := \alpha_i(x(t, s))$, we get

$$\int_{\Phi} d_E \alpha = \iint_{I \times [0,1]} (b^j \partial_t \tilde{\alpha}_j - a^j \partial_s \tilde{\alpha}_j - \tilde{\alpha}_i c_{jk}^i a^j b^k) dt ds.$$

Similarly,

$$\begin{aligned} \int_{\partial\Phi} \alpha &= \int_I (\tilde{\alpha}_i(t, 0) a^i(t, 0) - \tilde{\alpha}_i(t, 1) a^i(t, 1)) dt \\ &\quad - \int_{[0,1]} (\tilde{\alpha}_i(0, s) b^i(0, s) - \tilde{\alpha}_i(1, s) b^i(1, s)) ds. \end{aligned}$$

As we see, (3.3) holds for all $\tilde{\alpha}_i$ if and only if

$$\int_{\Phi} d_E \alpha = \int_{\partial\Phi} \alpha, \tag{3.4}$$

which can be understood as a generalized Stokes formula.

Remark 3.4. In fact, (3.3) is more general than (3.4) since $\tilde{\alpha}_i$ being the pull-back of α via the map Φ cannot be an arbitrary function of t and s . We can, however, easily overcome this drawback by using the graph of Φ in $\mathbb{T}\mathbb{R}|_I \times \mathbb{T}\mathbb{R}|_{[0,1]} \times E$ and $(\mathbb{T}\mathbb{R}|_I \times \mathbb{T}\mathbb{R}|_{[0,1]} \times E)$ -1-forms instead of Φ and E -1-forms.

Remark 3.5. The condition (3.1a) for a (and analogously (3.1b) for b) can be expressed in the Stokes-like way as well. Consider, namely, the map $\Phi_s(\cdot) := \Phi(\cdot, \theta_s) : \mathbb{T}\mathbb{R}|_I \longrightarrow E$. The admissibility of a reads as

$$\int_{\Phi_s} d_E f = \int_{\partial \Phi_s} f$$

for every $f \in C^\infty(M)$ and $s \in [0, 1]$.

E -homotopies on an integrable algebroid

Now we show that admissible paths and algebroid homotopies on an integrable algebroid $\mathcal{A}(\mathcal{G})$ are true paths and true homotopies on an integrating groupoid \mathcal{G} reduced to $\mathcal{A}(\mathcal{G})$ by means of the reduction map (A.1). We are using the language of Lie groupoids, so the reader unfamiliar with this topic should consult Appendix A.1.

Theorem 3.6 (integration). *Let $A(\mathcal{G}) \rightarrow M$ be a Lie algebroid of a Lie groupoid \mathcal{G} . Fix $x_0, y_0 \in M$ and an element $g_0 \in \alpha^{-1}(y_0) \cap \beta^{-1}(x_0)$.*

There is a 1-1 correspondence between:

- *bounded measurable admissible paths $a : [t_0, t_1] \longrightarrow \mathcal{A}(\mathcal{G})$ over an ACB path $x : [t_0, t_1] \longrightarrow M$ such that $x(t_0) = x_0$, and*
- *ACB paths $g : [t_0, t_1] \longrightarrow \mathcal{G}_{y_0}$ such that $g(t_0) = g_0$ and $x(t) = \beta(g(t))$.*

The correspondence is given by means of the reduction map (A.1); i.e., $a(t) = \mathcal{R}(\partial_t g(t)) = \text{TR}_{g(t)^{-1}}(\partial_t g(t))$.

Similarly, there is a 1-1 correspondence between:

- *bounded measurable algebroid homotopies $a, b : [t_0, t_1] \times [0, 1] \longrightarrow \mathcal{A}(\mathcal{G})$ over an ACB map $x : [t_0, t_1] \times [0, 1] \longrightarrow M$ such that $x(t_0, 0) = x_0$, and*
- *ACB homotopies $h : [t_0, t_1] \times [0, 1] \longrightarrow \mathcal{G}_{y_0}$ (i.e., h is ACB w.r.t. both variables) such that $h(t_0, 0) = g_0$ and $x(t, s) = \beta(h(t, s))$.*

Again, the correspondence is given by means of the reduction map (A.1); i.e., $a(t, s) = \mathcal{R}(\partial_t h(t, s)) = \text{TR}_{h(t,s)^{-1}}(\partial_t h(t, s))$ and $b(t, s) = \mathcal{R}(\partial_s h(t, s)) = \text{TR}_{h(t,s)^{-1}}(\partial_s h(t, s))$.

Proof. In the smooth case the result can be easily derived from Theorem A.2 and Corollary A.4. We know that (cf. considerations on page 87) smooth admissible paths and smooth algebroid homotopies in $\mathcal{A}(\mathcal{G})$ correspond to algebroid morphisms $\mathbb{T}\mathbb{R}|_{[t_0, t_1]} \longrightarrow \mathcal{A}(\mathcal{G})$ and $\mathbb{T}\mathbb{R}|_{[t_0, t_1]} \times \mathbb{T}\mathbb{R}|_{[0, 1]} \longrightarrow \mathcal{A}(\mathcal{G})$. The underlying manifolds are simply connected, hence the

morphisms can be lifted to smooth maps $g : [t_0, t_1] \rightarrow \mathcal{G}_{y_0}$ and $h : [t_0, t_1] \times [0, 1] \rightarrow \mathcal{G}_{y_0}$, which are unique up to the choice of the initial points (cf. Corollary A.4).

In the measurable case, however, the argument needs a little more attention. We will work in local coordinates (z^i) on \mathcal{G}_{y_0} , (x^a) on M , and linear coordinates (x^a, y^i) on $\mathcal{A}(\mathcal{G})$. We have induced coordinates (z^i, \dot{z}^j) on $T\mathcal{G}_{y_0}$ and (x^a, \dot{x}^b) on TM .

For $g \in \mathcal{G}_{y_0}$, $TR_{g^{-1}}$ maps $T_g\mathcal{G}_{y_0} = T_g^\alpha\mathcal{G}$ isomorphically into $\mathcal{A}(\mathcal{G})_{\beta(g)}$. In coordinates, $TR_{g^{-1}} : (z^i, \dot{z}^j) \mapsto (x^a, y^i)$ can be expressed as

$$\begin{aligned} x^a &= \beta^a(z), \\ y^i &= F_j^i(z) \dot{z}^j, \end{aligned}$$

where $\beta^a(z)$ and $F_j^i(z)$ are smooth and $F_j^i(z)$ is invertible. By $f_i^j(z)$ we will denote the inverse matrix of $F_j^i(z)$. The structure functions of the algebroid $\mathcal{A}(\mathcal{G})$ in these coordinates satisfy

$$\begin{aligned} \rho_i^a(\beta(z)) F_j^i(z) \dot{z}^j &= \frac{\partial \beta^a(z)}{\partial z^j} \dot{z}^j, \\ c_{jk}^i(\beta(z)) F_m^j(z) F_n^k(z) \dot{z}^m \dot{z}^n &= \left(\frac{\partial F_n^i(z)}{\partial z_m} - \frac{\partial F_m^i(z)}{\partial z_n} \right) \dot{z}^m \dot{z}^n, \end{aligned}$$

since ρ is the reduced $T\beta$, and the $\mathcal{A}(\mathcal{G})$ -bracket is the reduced Lie bracket on \mathcal{G} . From the above we get

$$\begin{aligned} \rho_i^a(\beta(z)) &= \frac{\partial \beta^a(z)}{\partial z^j} f_i^j(z), \\ c_{jk}^i(\beta(z)) &= \left(\frac{\partial F_n^i(z)}{\partial z_m} - \frac{\partial F_m^i(z)}{\partial z_n} \right) f_j^m(z) f_k^n(z). \end{aligned}$$

To prove the first part of the assertion, observe that, if $g : [t_0, t_1] \rightarrow \mathcal{G}_{y_0}$ is an ACB path over an ACB path $x : [t_0, t_1] \rightarrow M$, then the derivative $\partial_t g(t) \in T_{g(t)}\mathcal{G}_{y_0}$ is a bounded measurable path, and so is $a(t) = \mathcal{R}(\partial_t g(t)) = TR_{g(t)^{-1}}(\partial_t g(t))$, since \mathcal{R} is smooth. Clearly, $\dot{x}(t) = T\beta(\partial_t g(t)) = \rho(a(t))$ (cf. diagram (A.1)), so $a(t)$ is a bounded measurable $\mathcal{A}(\mathcal{G})$ -path.

Conversely, consider a bounded measurable admissible path $a : [t_0, t_1] \rightarrow \mathcal{A}(\mathcal{G})$ over an ACB path $x : [t_0, t_1] \rightarrow M$. For every $t \in [t_0, t_1]$ and all g satisfying $\beta(g) = x(t)$ we may lift $a(t) \in \mathcal{A}(\mathcal{G})_{x(t)}$ to a vector $A(t, g) := TR_g(a(t)) \in T_g\mathcal{G}_{y_0}$. We would like to define $g(t)$ as a solution of the differential equation in \mathcal{G}_{y_0}

$$\partial_t g(t) = A(g, t)$$

with the initial condition $g(t_0) = g_0$. Then, clearly, $TR_{g(t)^{-1}}(\partial_t g(t)) = TR_{g(t)^{-1}}A(g(t), t) = a(t)$ as in the assertion. The problem is that, since $A(g, t)$ is defined only on a subset of \mathcal{G}_{y_0} it is not clear that the solution exists, nor that it is unique. To overcome this difficulty consider a differential equation on \mathcal{G}_{y_0} given in local coordinates by

$$\dot{z}^i = f_j^i(z) a^j(t). \quad (3.5)$$

It satisfies the assumptions of Theorem B.12 for measurable ODEs, so it has an ACB solution $z(t)$, unique up to the choice of the initial point. In particular, let $z(t)$ be the solution with $z(t_0) = g_0 \in \mathcal{G}_{y_0}$. The base trajectory $\tilde{x}(t) = \beta(z(t))$ satisfies

$$\begin{aligned}\partial_t \tilde{x}^a(t) &= \frac{\partial \beta^a(z(t))}{\partial z^j} \dot{z}^j(t) = \frac{\partial \beta^a(z(t))}{\partial z^j} f_j^i(z(t)) a^i(t) = \rho_i^a(\tilde{x}(t)) a^i(t), \\ \tilde{x}(t_0) &= \beta(z(t_0)) = \beta(g_0) = x_0.\end{aligned}$$

On the other hand, by admissibility of $a(t)$, we have $\partial_t x^a(t) = \rho_i^a(x(t)) a^i(t)$ and $x(t_0) = x_0$, hence; clearly, $\tilde{x}(t) = x(t)$. This, in turn, implies that $\dot{z}^i(t) = f_j^i(z(t)) a^j(t) = A^i(z(t), t)$, i.e., $g(t) = z(t)$ as above is well defined and unique.

Now consider a homotopy $h : [t_0, t_1] \times [0, 1] =: K \longrightarrow \mathcal{G}_{y_0}$ over $x : [t_0, t_1] \times [0, 1] \longrightarrow M$, which is ACB w.r.t. both variables. In local coordinates it is given by $z^i(s, t)$. Repeating the argument from the previous part, we can prove that the maps $t \mapsto a(t, s) := \text{TR}_{h(t,s)^{-1}}(\partial_t h(t, s))$ and $s \mapsto b(t, s) := \text{TR}_{h(t,s)^{-1}}(\partial_s h(t, s))$ are bounded measurable admissible paths over $t \mapsto x(t, s)$ and $s \mapsto x(t, s)$, respectively. In local coordinates,

$$\begin{aligned}a^i(t, s) &= F_j^i(z(t, s)) A^j(t, s), \\ b^i(t, s) &= F_j^i(z(t, s)) B^j(t, s),\end{aligned}$$

where we denoted $A^i(t, s) := \partial_t z^i(t, s)$ and $B^i(t, s) = \partial_s z^i(t, s)$.

Since h is a homotopy, we have

$$\iint_K z^i(t, s) \partial_t \partial_s \phi_i(t, s) dt ds = \iint_K z^i(t, s) \partial_s \partial_t \phi_i(t, s) dt ds,$$

for every $\phi_i \in C^\infty(K)$. Integrating the above equality several times by parts, we get that $A^i(t, s)$ and $B^i(t, s)$ satisfy the differential equation

$$\partial_s A^i(t, s) \stackrel{\text{WT}}{=} \partial_t B^i(t, s).$$

Now calculating the WT-derivatives of $a^i(t, s)$ and $b^i(t, s)$ we get

$$\begin{aligned}\partial_t b^i(t, s) - \partial_s a^i(t, s) &\stackrel{\text{WT}}{=} \partial_t (F_j^i(z(t, s)) B^j(t, s)) - \partial_s (F_j^i(z(t, s)) A^j(t, s)) \\ &\stackrel{\text{WT}}{=} \left(\frac{\partial F_j^i}{\partial z^m}(z(t, s)) - \frac{\partial F_j^i}{\partial z^m}(z(t, s)) \right) B^j(t, s) A^m(t, s) \\ &\quad + F_j^i(z(t, s)) (\partial_t B^j(t, s) - \partial_s A^j(t, s)) \\ &= c_{jk}^i(\beta(z(t, s))) b^j(t, s) a^k(t, s) + 0 = c_{jk}^i(x(t, s)) b^j(t, s) a^k(t, s).\end{aligned}$$

We see that (a, b) is an $\mathcal{A}(\mathcal{G})$ -homotopy.

Conversely, let $a, b : [t_0, t_1] \times [0, 1] \longrightarrow \mathcal{A}(\mathcal{G})$ over $x : [t_0, t_1] \times [0, 1] \longrightarrow M$ be an algebroid homotopy. By the first part of the assertion we can uniquely integrate the admissible path $s \mapsto b_0(t_0, s)$ to an ACB path $g_0(s) \in \mathcal{G}_{y_0}$ with $g_0(0) = g_0$. Next we can uniquely integrate each admissible path $t \mapsto a(t, s)$ to an ACB path $g(t, s) \in \mathcal{G}_{y_0}$ such that

$g(t_0, s) = g_0(s)$. In local coordinates $g(t, s)$ is a solution of the differential equation (cf. the previous part of this proof)

$$\begin{aligned}\partial_t z^i(t, s) &= f_j^i(z(t, s))a^j(t, s), \\ z^i(t_0, s) &= z_0^i(s),\end{aligned}$$

where $z_0(s) = g_0(s)$ is ACB. By Theorem B.13, $g(t, s)$ is ACB w.r.t. both variables. Now $g(t, s)$ is an ACB homotopy in \mathcal{G}_{y_0} hence, as has already been proved, it reduces to an algebroid homotopy $\tilde{a}, \tilde{b} : [t_0, t_1] \rightarrow \mathcal{A}(\mathcal{G})$. By construction, $\tilde{a}(t, s) = a(t, s)$ and $\tilde{b}(t_0, s) = b(t_0, s)$. We see that (a, b) and (\tilde{a}, \tilde{b}) are two WT-solutions of (3.2) with the same initial-point $\mathcal{A}(\mathcal{G})$ -homotopy $b(t_0, s)$. By Lemma 3.3 $b(t, s) = \tilde{b}(t, s)$. \square

Remark 3.7. The above theorem is closely related to the ideas of [Crainic & Fernandes, 2003]. The correspondence between algebroid homotopies and homotopies in an integrating groupoid may be used to address the question about integrability of Lie algebroids (see remark on page 86).

In fact, we can also use it to prove Theorem A.2 of Mackenzie and Xu. To sketch the idea, let us concentrate on the case when $\mathcal{H} = S \times S$ is a pair groupoid, with S simply connected. Consider a morphism of Lie algebroids $\Phi : \text{TS} \rightarrow \mathcal{A}(\mathcal{G})$ over $f : S \rightarrow M$. Fixing points $x_0 \in S$ and $g_0 \in \mathcal{G}_{f(x_0)} \cap \beta^{-1}(f(x_0))$, we can attach to each sufficiently regular curve $\gamma : [0, 1] \rightarrow S$, originated at $\gamma(0) = x_0$, a curve $g : [0, 1] \rightarrow \mathcal{G}_{f(x_0)}$, with $g(0) = g_0$, being the lift of an admissible curve $\Phi(\dot{\gamma}(\cdot)) : [0, 1] \rightarrow \mathcal{A}(\mathcal{G})$. Now, if γ_0 and γ_1 are two curves such that $\gamma_0(0) = \gamma_1(0) = x_0$ and $\gamma_0(1) = \gamma_1(1)$, then, since S is simply connected, there exists a homotopy $\gamma(t, s)$ in S (with fixed end-points) joining $\gamma(\cdot, 0) = \gamma_0(\cdot)$ and $\gamma(\cdot, 1) = \gamma_1(\cdot)$. The lift of the $\mathcal{A}(\mathcal{G})$ -homotopy $(\Phi(\partial_t \gamma(t, s)), \Phi(\partial_s \gamma(t, s)))$ is a homotopy $g(t, s)$ in $\mathcal{G}_{f(x_0)}$ (with fixed end-points) joining the lifts of $\gamma_0(\cdot)$ and $\gamma_1(\cdot)$. Consequently, the map $\tilde{\Phi} : \gamma(1) \mapsto g(1), S \rightarrow \mathcal{G}_{f(x_0)}$ is well defined. One can prove that $\mathcal{R} \circ \text{T}\tilde{\Phi} = \Phi$. The presence of such a map is equivalent to the integrability of Φ (see Corollary A.4). A similar argument can be used to prove Theorem A.2 in full generality.

Corollary 3.8. *Theorem 3.6 establishes the equivalence between $\mathcal{A}(\mathcal{G})$ -paths/homotopies and standard paths/homotopies in a single α -fibre in the groupoid \mathcal{G} . For the groupoid $\mathcal{G}_P = P \times P/G$ and the associated Atiyah algebroid TP/G , these fibres are canonically isomorphic to P , so the E -homotopies are just standard homotopies in P reduced to TP/G . The two are equivalent up to the choice of the initial point.*

E -homotopy classes

Definition 3.9. Two measurable E -paths $a_0, a_1 : [t_0, t_1] \rightarrow E$ are *E -homotopic* iff there exists an E -homotopy (a, b) with fixed end-points (i.e., $b(t_0, \cdot) \equiv 0 \equiv b(t_1, \cdot)$) between a_0 and a_1 . Being E -homotopic is an equivalence relation.

An equivalence class of an element a will be denoted by $[a]$ (or sometimes $[a(t)]_{t \in [t_0, t_1]}$) and called an *E -homotopy class*.

So far, the above definition does not allow us to compare the E -homotopy classes of E -paths defined on different time intervals. Therefore, we will add a natural condition that a composition with a null path not change the equivalence class: $[a \circ \theta_{x(t_1)}] = [a] = [\theta_{x(t_0)} \circ a]$.

Observe that, since two E -homotopies with fixed end-points are composable (iff the final base point of the first coincides with the initial base point of the second), the composition of E -homotopies defines a multiplication of E -homotopy classes by a natural formula

$$[a] \cdot [\bar{a}] := [a \circ \bar{a}] \quad \text{when } a \text{ and } \bar{a} \text{ are composable.}$$

Consider now an algebroid E and two smooth algebroid morphisms $\Phi_0 : TS_0 \rightarrow E$ and $\Phi_1 : TS_1 \rightarrow E$ over $\phi_0 : S_0 \rightarrow M$ and $\phi_1 : S_1 \rightarrow M$, respectively.

Definition 3.10. We say that measurable E -paths a_0 and a_1 are *E -homotopic relative to the morphisms Φ_0 and Φ_1* iff there exists an E -homotopy (a, b) between a_0 and a_1 , and AC paths $z : [0, 1] \rightarrow S_0$ and $w : [0, 1] \rightarrow S_1$, such that $b(t_0, s) = \Phi_0(\partial_s z(s))$ and $b(t_1, s) = \Phi_1(\partial_s w(s))$. In other words, initial-point E -homotopies lie in $\text{Im } \Phi_0$ and $\text{Im } \Phi_1$, respectively. Note that $b(t_0, \cdot)$ and $b(t_1, \cdot)$ are admissible as images of admissible paths under an algebroid morphism.

The relation of being relatively E -homotopic is again an equivalence relation and we may again speak of the equivalence classes (*relative E -homotopy classes*). A class of an element a will be denoted by $[a] \bmod(\Phi_0, \Phi_1)$.

Remark 3.11. E -paths a_0 and a_1 are E -homotopic iff they are E -homotopic relative to a morphism ι_{x_0} , which maps $S_0 = \{\text{pt}\} \approx TS_0$ to a null vector θ_{x_0} , where $x_0 = \tau \circ a_0(t_0) = \tau \circ a_1(t_0)$, and a morphism ι_{x_1} defined analogously for $x_1 = \tau \circ a_0(t_1) = \tau \circ a_1(t_1)$.

The interpretation of E -homotopy classes

In light of Theorem 3.6, an interpretation of the notion of E -homotopy classes is clear. If $E = \mathcal{A}(\mathcal{G})$ is an integrable Lie algebroid and $(a, b) : [t_0, t_1] \times [0, 1] \rightarrow E$ is an algebroid homotopy with fixed end points, we can lift it to the true homotopy $g(t, s)$ in a single α -fibre \mathcal{G}_{y_0} of \mathcal{G} . The E -paths $a(\cdot, 0)$ and $a(\cdot, 1)$ correspond to $g(\cdot, 0)$ and $g(\cdot, 1)$, respectively, and null paths $b(t_0, \cdot)$ and $b(t_1, \cdot)$ to constant paths $g(t_0, \cdot)$ and $g(t_1, \cdot)$, respectively. In other words, $g(t, s)$ is a homotopy in \mathcal{G}_{y_0} between $g(\cdot, 0)$ and $g(\cdot, 1)$ with fixed end-points. Consequently, we can interpret E -homotopy classes as reduced homotopy classes from an α -fibre \mathcal{G}_{y_0} of a Lie groupoid \mathcal{G} to the associated Lie algebroid $E = \mathcal{A}(\mathcal{G})$. In particular, if \mathcal{G}_P is a gauge groupoid of a principal G -bundle $G \rightarrow P \rightarrow M$, all α -fibres are isomorphic to P (cf. Appendix A.1), hence algebroid homotopy classes in the Atiyah algebroid TP/G are the standard homotopy classes in P reduced to TP/G by the G -action.

For relative E -homotopy classes things are a little more complicated. Assume that $E = \mathcal{A}(\mathcal{G})$ is an integrable Lie algebroid and $a, b : [t_0, t_1] \times [0, 1] \rightarrow \mathcal{A}(\mathcal{G})$ is an algebroid homotopy relative to (Φ_0, Φ_1) . Let $z(\cdot) \subset S_0$ and $w(\cdot) \subset S_1$ be as in the Definition 3.10. Assume, in addition, that Φ_0 and Φ_1 are integrable. By Theorem 3.6 we can lift (a, b) to the homotopy $g(t, s) \in \mathcal{G}_{y_0} \subset \mathcal{G}$. By Corollary A.4 we can lift Φ_0 to a smooth map $\tilde{\Phi}_0 : S_0 \rightarrow \mathcal{G}_{y_0}$ such that $\tilde{\Phi}_0(z(0)) = g(t_0, 0)$, and we can lift Φ_1 to a smooth map

$\tilde{\Phi}_1 : S_1 \rightarrow \mathcal{G}_{y_0}$ such that $\tilde{\Phi}_1(w(0)) = g(t_1, 0)$. Paths $\tilde{\Phi}_0(z(\cdot))$ and $g(t_0, \cdot)$ correspond to the same $\mathcal{A}(\mathcal{G})$ -path $b(t_0, \cdot)$ and have the same initial point, hence are equal. Similarly, $\tilde{\Phi}_1(w(\cdot)) = g(t_1, \cdot)$. In other words, $g(t, s)$ is a homotopy in \mathcal{G}_{y_0} joining $g(\cdot, 0)$ and $g(\cdot, 1)$ with end-points in the images of $\tilde{\Phi}_0$ and $\tilde{\Phi}_1$.

To sum up, (a, b) is a reduction of a homotopy in \mathcal{G}_{y_0} with end-points contained in the images of $\tilde{\Phi}_0$ and $\tilde{\Phi}_1$ integrating Φ_0 and Φ_1 .

Note that algebroid morphisms Φ_0 and Φ_1 need not to be integrable. In such a case the interpretation given above is still valid, but locally. Let namely $U_0 \subset S_0$ be an open simply connected neighbourhood of $z(0)$ and let $U_1 \subset S_1$ be an open simply connected neighbourhood of $w(0)$. Now, by Corollary A.4, we can lift $\Phi_0|_{U_0} : TU_0 \rightarrow \mathcal{A}(\mathcal{G})$ and $\Phi_1|_{U_1} : TU_0 \rightarrow \mathcal{A}(\mathcal{G})$ to $\tilde{\Phi}_0|_{U_0} : U_0 \rightarrow \mathcal{G}_{y_0}$ and $\tilde{\Phi}_1|_{U_1} : U_1 \rightarrow \mathcal{G}_{y_0}$, respectively. Hence, if $b(t_0, \cdot) \subset \text{Im } \Phi_0|_{U_0}$ and $b(t_1, \cdot) \subset \text{Im } \Phi_1|_{U_1}$ we can still interpret the relative $\mathcal{A}(\mathcal{G})$ -homotopy (a, b) as the homotopy in \mathcal{G}_{y_0} with end-points in the images of $\tilde{\Phi}_0|_{U_0}$ and $\tilde{\Phi}_1|_{U_1}$, reduced to the algebroid $\mathcal{A}(\mathcal{G})$.

Another, more universal approach is the following. Consider the universal covers $\pi_0 : (\tilde{S}_0, \tilde{z}(0)) \rightarrow (S_0, z(0))$ and $\pi_1 : (\tilde{S}_1, \tilde{w}(0)) \rightarrow (S_1, w(0))$. Now take Lie algebroid morphisms $\Phi'_i := \Phi_i \circ \pi_i : T\tilde{S}_i \rightarrow \mathcal{A}(\mathcal{G})$, for $i = 0, 1$. They are clearly integrable to $\tilde{\Phi}'_i : \tilde{S}_i \rightarrow \mathcal{G}_{y_0}$, since \tilde{S}_0 and \tilde{S}_1 are simply connected. As $b(t_0, \cdot) \subset \text{Im } \Phi'_0$ and $b(t_1, \cdot) \subset \text{Im } \Phi'_1$, we can interpret (a, b) as a reduced homotopy in \mathcal{G}_{y_0} with end-points in the images of $\tilde{\Phi}'_0$ and $\tilde{\Phi}'_1$.

In particular, if $\mathcal{G} = \mathcal{G}_P$ is a gauge groupoid of a principal G -bundle $G \rightarrow P \rightarrow M$, then $\mathcal{G}_{y_0} \approx P$. Relative TP/G -homotopies are homotopies in P with end-points in the images of maps $\tilde{\Phi}'_0 : \tilde{S}_0 \rightarrow P$ and $\tilde{\Phi}'_1 : \tilde{S}_1 \rightarrow P$, reduced by the G -action.

3.2 Fundamental properties of E -homotopies

E -homotopies as families of E -paths

The following lemma emphasizes the role of AL algebroids. Roughly speaking, it turns out that for AL algebroids one-parameter families of E -paths are E -homotopies.

Lemma 3.12 (generating E -homotopies). *Let E be an AL algebroid, and let $a : I \times [0, 1] \rightarrow E$ be a one-parameter family of bounded measurable E -paths (that is, $t \mapsto a(t, s)$ is admissible for every s) covering $x : I \times [0, 1] \rightarrow M$. Assume that $a(t, s)$ is ACB w.r.t. s ; that is, $\partial_s a(t, s)$ is defined a.e. and is bounded measurable w.r.t. both variables. Let $b_0(s)$ be an arbitrary bounded measurable E -path covering $x(t_0, s)$.*

Then there exists a unique E -homotopy $a, b : I \times [0, 1] \rightarrow E$ such that $b(t_0, s) = b_0(s)$. Moreover, $b(t, s)$ is ACB w.r.t. t (that is, $\partial_t b(t, s)$ is defined a.e. and is bounded measurable w.r.t. both variables).

Proof. By the definition of an E -homotopy, $b(t, s)$ should be a map covering $x(t, s)$ such that $s \mapsto b(t, s)$ is admissible and (3.3) is WT-satisfied. Observe that, since $\partial_s a(t, s)$ is

well defined a.e., the system of equations

$$\partial_t b^i(t, s) = \partial_s a^i(t, s) + c_{jk}^i(x(t, s)) b^j(t, s) a^k(t, s) \quad (3.6)$$

for $b(t, s)$ satisfies the assumptions of Theorem B.13. Consequently, it has a unique Carathéodory solution $b^i(t, s)$ for a given initial condition $b^i(t_0, s) = b_0^i(s)$. The solution $b(t, s)$ is ACB w.r.t. t and, since the parameter- s -dependence of both right-hand side of (3.6) and the initial condition is bounded measurable, so is the s -dependence of the solution $b(t, s)$. Consequently, the right-hand side of (3.6) is bounded and measurable w.r.t. both t and s , and hence so is $\partial_t b(t, s)$ —the left-hand side of (3.6). Clearly, a and b are regular enough to satisfy the assumptions of Theorem B.18, so the integral condition (3.3) holds.

To prove that thus constructed (a, b) is indeed an E -homotopy, it is enough to show that $s \mapsto b(t, s)$ is admissible for every fixed t . Consider a map

$$\chi^a(t, s) := \partial_s x^a(t, s) - \rho_i^a(x(t, s)) b^i(t, s).$$

We shall show that $\chi^a = 0$ a.e. Observe that, since $a(t, s)$ is a family of admissible paths, we have

$$\partial_t x^a(t, s) = \rho_k^a(x(t, s)) a^k(t, s) \text{ a.e.} \quad (3.7)$$

The right-hand side of this equation is differentiable with respect to s , and hence so is the left-hand side, and

$$\partial_s \partial_t x^a(t, s) = \frac{\partial \rho_k^a}{\partial x^b}(x(t, s)) \partial_s x^b(t, s) a^k(t, s) + \rho_k^a(x(t, s)) \partial_s a^k(t, s).$$

Consequently, as $\partial_t \partial_s x^a(t, s) \stackrel{\text{WT}}{=} \partial_s \partial_t x^a(t, s)$ (since $x(t, s)$ is a true homotopy in M),

$$\begin{aligned} \partial_t \chi^a &= \partial_t \partial_s x^a - \partial_t (\rho_k^a b^k) \stackrel{\text{WT}}{=} \partial_s \partial_t x^a - \partial_t (\rho_k^a b^k) \\ &= \frac{\partial \rho_k^a}{\partial x^b} \partial_s x^b a^k + \rho_k^a \partial_s a^k - \frac{\partial \rho_i^a}{\partial x^b} \partial_t x^b b^i - \rho_i^a \partial_t b^i, \end{aligned}$$

which, in view of (3.7) and (3.6), equals

$$\begin{aligned} &\frac{\partial \rho_k^a}{\partial x^b} \partial_s x^b a^k + \rho_k^a \partial_s a^k - \frac{\partial \rho_i^a}{\partial x^b} \rho_k^b a^k b^i - \rho_i^a (\partial_s a^i + c_{jk}^i b^j a^k) = \\ &= \left(\frac{\partial}{\partial x^b} \rho_k^a \right) a^k \chi^b + \left[\left(\frac{\partial}{\partial x^b} \rho_k^a \right) \rho_j^b - \left(\frac{\partial}{\partial x^b} \rho_j^a \right) \rho_k^b - \rho_i^a c_{jk}^i \right] b^j a^k. \end{aligned}$$

Since E is an AL algebroid, the last term vanishes and we have

$$\partial_t \chi^a \stackrel{\text{WT}}{=} \left(\frac{\partial}{\partial x^b} \rho_k^a \right) a^k \chi^b.$$

Thus χ^a is a WT-solution of a linear differential equation with measurable r.h.s. and the initial condition $\chi^a(t_0, s) = 0$ (since $b_0(s)$ is admissible). Repeating the argument from the proof of Lemma B.19 we conclude that $\chi^a = 0$ a.e. \square

It turned out that, when a skew-algebroid is almost Lie, E -homotopies are the true homotopies in the space of E -paths (i.e. one-parameter families of E -paths). This has already been observed in [Grabowska & Grabowski, 2008, Thm. 3] in a slightly different form.

Relation between E -homotopies with and without fixed end-points, reparametrisation

Lemma 3.13. *Let $a : [t_0, t_1] \rightarrow E$ be a bounded measurable E -path over $x(t)$, and let $h : [0, 1] \rightarrow [t_0, t_1]$ be an invertible C^1 -function. Define*

$$\begin{aligned} a(t, s) &:= \frac{h(s) - t_0}{t_1 - t_0} a \left(t_0 + \frac{t - t_0}{t_1 - t_0} (h(s) - t_0) \right), \\ b(t, s) &:= \frac{t - t_0}{t_1 - t_0} \dot{h}(s) a \left(t_0 + \frac{t - t_0}{t_1 - t_0} (h(s) - t_0) \right). \end{aligned} \quad (3.8)$$

Then the pair (a, b) is an E -homotopy over $x(t, s) := x \left(t_0 + \frac{t - t_0}{t_1 - t_0} (h(s) - t_0) \right)$.

Proof. For notation simplicity assume that $[t_0, t_1] = [0, 1]$. Then

$$\begin{aligned} a(t, s) &= h(s)a(th(s)) \quad \text{and} \\ b(t, s) &= th(s)a(th(s)). \end{aligned}$$

First, note that $t \mapsto a(t, s)$ and $s \mapsto b(t, s)$ are admissible. Indeed, from $\dot{x}(t) = \rho(a(t))$ we deduce that

$$\partial_t x(t, s) = h(s)\dot{x}(th(s)) = h(s)\rho(a(th(s))) = \rho(a(t, s)).$$

Similarly, $\partial_s x(t, s) = \rho(b(t, s))$. Now we will check that

$$\partial_t [th(s)a^i(th(s))] \stackrel{W}{=} \partial_s [h(s)a^i(th(s))] + c_{jk}^i(x(t, s)) [th(s)a^j(th(s))] \cdot [h(s)sa^k(th(s))].$$

By the skew-symmetry of c_{jk}^i , the last term vanishes, so we have to check if

$$\partial_t [th(s)a^i(th(s))] \stackrel{W}{=} \partial_s [h(s)a^i(th(s))].$$

The latter is certainly true, as both sides are equal $\dot{h}(s)a^i(th(s)) + th(s)\dot{h}(s)A^i(th(s))$, where $A^i(t)$ is the distributive derivative of $a^i(t)$.

To finish the proof we shall show that $a(t, s)$ and $b(t, s)$ satisfy the regularity conditions (B.10) and (B.11).

Assume that $h(0) = t_0 = 0$ and $h(1) = t_1 = 1$ (the case $h(1) = t_0 = 0$ and $h(0) = t_1 = 1$ is completely analogous). Now

$$\int_0^1 \int_0^\varepsilon |a(t, s) - a(t, 0)| \frac{1}{\varepsilon} ds dt = \frac{1}{\varepsilon} \int_0^\varepsilon \int_0^1 |h(s)a(th(s))| dt ds \leq \frac{1}{\varepsilon} \int_0^\varepsilon h(s) \|a\| ds.$$

The later converges to 0 as $\varepsilon \rightarrow 0$, since $h(s) \xrightarrow{s \rightarrow 0} h(0) = 0$. Next,

$$\begin{aligned} & \int_0^1 \int_{1-\varepsilon}^1 |a(t, s) - a(t, 1)| \frac{1}{\varepsilon} ds dt = \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 \int_0^1 |h(s)a(th(s)) - h(1)a(th(1))| dt ds \\ & \leq \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 h(s) \int_0^1 |a(th(s)) - a(th(1))| dt ds + \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 |h(s) - h(1)| \int_0^1 |a(th(s))| dt ds. \end{aligned}$$

The second factor converges to 0 as $\varepsilon \rightarrow 0$ because $h(s)$ is continuous at $s = 1$ and a is bounded. By Lemma B.4 the measurable function $g(s) := \int_0^1 |a(th(s)) - a(th(1))| dt$ is regular at $s = 0$ and, moreover, $g(s) = 0$. We conclude that

$$\frac{1}{\varepsilon} \int_{1-\varepsilon}^1 h(s)g(s)ds \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0.$$

Consequently, $\int_0^1 \int_{1-\varepsilon}^1 |a(t, s) - a(t, 1)| \frac{1}{\varepsilon} ds dt \xrightarrow{\varepsilon \rightarrow 0} 0$ and conditions (B.10) are fulfilled.

Now check (B.11). The first of the two conditions is a matter of a simple estimation:

$$\int_0^1 \int_0^\varepsilon |b(t, s) - b(0, s)| \frac{1}{\varepsilon} dt ds = \frac{1}{\varepsilon} \int_0^\varepsilon \int_0^1 |t\dot{h}(s)a(th(s))| ds dt \leq \frac{1}{\varepsilon} \int_0^\varepsilon t \|\dot{h}\| \cdot \|a\| dt \xrightarrow{\varepsilon \rightarrow 0} 0.$$

For the second we can estimate:

$$\begin{aligned} \int_0^1 \int_{1-\varepsilon}^1 |b(t, s) - b(1, s)| \frac{1}{\varepsilon} dt ds &= \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 \int_0^1 |t\dot{h}(s)a(th(s)) - \dot{h}(s)a(h(s))| ds dt \\ &\leq \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 t \|\dot{h}\| \int_0^1 |a(th(s)) - a(h(s))| ds dt + \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 |1-t| \cdot \|\dot{h}\| \int_0^1 |a(h(s))| ds dt. \end{aligned}$$

The last factor clearly converges to 0 as $\varepsilon \rightarrow 0$. Using Lemma B.4 we show that the measurable function $k(t) := \int_0^1 |a(th(s)) - a(h(s))| ds$ is regular at $t = 1$ and, moreover, $k(1) = 0$. We conclude that

$$\frac{1}{\varepsilon} \int_{1-\varepsilon}^1 t \|\dot{h}\| k(t) dt \xrightarrow{\varepsilon \rightarrow 0} 0,$$

which proves that conditions (B.11) are satisfied. By Theorem B.18 the pair (a, b) is a WT-solution of (3.2), and hence E -homotopy. \square

As a corollary we obtain the following fact.

Lemma 3.14 (shrinking an E -path). *Let $a : [0, 1] \rightarrow E$ be a measurable E -path over $x(t)$. Define $a(t, s) := sa(ts)$ and $b(t, s) := ta(ts)$ for $t, s \in [0, 1]$. The pair (a, b) is an E -homotopy over $x(t, s) = x(ts)$. Its initial-point E -homotopy is $\theta_{x(0)}$, and the final-point E -homotopy is a .*

Similarly, consider $\tilde{a}(t, s) := (1-s)a(1-(1-t)(1-s))$ and $\tilde{b}(t, s) := (1-t)a(1-(1-t)(1-s))$ where $t, s \in [0, 1]$. The pair (\tilde{a}, \tilde{b}) is an E -homotopy over $\tilde{x}(t, s) := x(1-(1-t)(1-s))$. Its initial-point E -homotopy is a , and the final-point E -homotopy is $\theta_{x(1)}$.

Proof. The assertion follows from Lemma 3.13. For (a, b) we simply take $a(t)$ and $h(s) = s$. For (\tilde{a}, \tilde{b}) we use Lemma 3.13 with $\tilde{a}(t) := a(1-t)$ defined on an interval $[t_0 = 1, t_1 = 0]$ and $h(s) = s$. \square

We can now state the following important result.

Lemma 3.15. *Let $a, b : [t_0, t_1] \times [0, 1] \longrightarrow E$ be an E -homotopy covering $x : [t_0, t_1] \times [0, 1] \longrightarrow M$. Then we have the following equality of E -homotopy classes:*

$$[a(t, 0)]_{t \in [t_0, t_1]} [b(t_1, s)]_{s \in [0, 1]} = [b(t_0, s)]_{s \in [0, 1]} [a(t, 1)]_{t \in [t_0, t_1]}.$$

Proof. The first part of Lemma 3.14, applied to the curve $s \mapsto b(t_0, s)$, gives us the existence of E -homotopy $c, d : [0, 1] \times [0, 1] \longrightarrow E$ such that $c(t, 0) = \theta_{x(t_0)}$, $c(s, 1) = b(t_0, s)$, $d(0, s) = \theta_{x(t_0)}$, and $d(1, s) = b(t_0, s)$. Similarly, using the second part of Lemma 3.14 for $s \mapsto b(t_1, s)$, we obtain E -homotopy $e, f : [0, 1] \times [0, 1] \longrightarrow E$ such that $e(s, 0) = b(t_1, s)$, $e(t, 1) = \theta_{x(t_1)}$, $f(0, s) = b(t_1, s)$, and $f(1, s) = \theta_{x(t_1)}$.

Clearly, E -homotopies (c, d) , (a, b) , and (e, f) are composable and their composition is an E -homotopy with fixed end-points which establishes an equivalence of E -paths $\theta_{x(t_0)} \circ a(\cdot, 0) \circ b(t_1, \cdot)$ and $b(t_0, \cdot) \circ a(\cdot, 1) \circ \theta_{x(t_1)}$. \square

Remark 3.16. The above lemma is very important, as it shows the relation between E -homotopies with and without fixed end-points. If $a, b : I \times [0, 1] \longrightarrow E$ is an E -homotopy joining a_0 and a_1 , then the composition of a_0 with the final-point E -homotopy $b(t_1, \cdot)$ is equivalent to the composition of the initial-point E -homotopy $b(t_0, \cdot)$ with a_1 . Thus, if the initial-point E -homotopy $b(t_0, \cdot)$ vanishes, in order to check whether $[a_0] = [a_1]$, it is enough to check whether $[b(t_1, \cdot)] = 0$. Thus the problem of equivalence of a_0 and a_1 can be solved by investigating the final-point E -homotopy. Similarly, we can address the problem of relative E -homotopy equivalence $[a_0] = [a_1] \bmod(\Phi_0, \Phi_1)$ by studying the classes $[b(t_0, \cdot)]$ and $[b(t_1, \cdot)]$.

Finally, as a corollary from Lemmas 3.13 and 3.15 we obtain a result about reparametrisation of E -paths.

Lemma 3.17 (reparametrization). *Let $a : [t_0, t_1] \longrightarrow E$ be a measurable E -path, and let $h : [0, 1] \rightarrow [t_0, t_1]$ be an invertible C^1 -function. Define $\tilde{a}(t) := \dot{h}(t)a(h(t))$ for $t \in [0, 1]$. Then*

$$[\tilde{a}(t)]_{t \in [0, 1]} = [a(t)]_{t \in [t_0, t_1]} \quad \text{if } h(0) = t_0 \text{ and } h(1) = t_1, \quad (3.9)$$

$$[a(t)]_{t \in [t_0, t_1]} \cdot [\tilde{a}(t)]_{t \in [0, 1]} = 0 \quad \text{if } h(0) = t_1 \text{ and } h(1) = t_0. \quad (3.10)$$

Proof. Consider a homotopy (3.8) from Lemma 3.13. If $h(0) = t_0$ and $h(1) = t_1$ we have $a(t, 0) = 0$, $a(t, 1) = a(t)$, $b(t_0, s) = 0$, and $b(t_1, s) = \dot{h}(s)a(h(s)) = \tilde{a}(s)$. By Lemma 3.15,

$$[\theta_{x(t_0)}] \cdot [\tilde{a}(s)]_{s \in [0, 1]} = [\theta_{x(t_0)}] \cdot [a(t)]_{t \in [t_0, t_1]}.$$

Analogously, for $h(0) = t_1$ and $h(1) = t_0$ we have $a(t, 0) = a(t)$, $a(t, 1) = 0$, $b(t_0, s) = 0$, and $b(t_1, s) = \dot{h}(s)a(h(s)) = \tilde{a}(s)$. From Lemma 3.15 we deduce that

$$[a(t)]_{t \in [t_0, t_1]} \cdot [\tilde{a}(s)]_{s \in [0, 1]} = [\theta_{x(t_0)}] \cdot [\theta_{x(t_0)}] = 0.$$

\square

Optimal control problems on AL algebroids

In this chapter we introduce the notion of a control system and an optimal control problem on algebroids. Much attention is paid to motivate these definitions. We show that equivariant control systems and optimal control problems on a Lie groupoid \mathcal{G} lead naturally to system and problems on the associated Lie algebroid $\mathcal{A}(\mathcal{G})$. What is more, our definitions coincide with the standard ones for special cases of tangent algebroid and Atiyah algebroid. At the end, we define the natural notion of algebroid homotopy associated with a control system.

From this chapter on, our attention is restricted to AL algebroids only. This choice is justified by the properties of algebroid homotopies on AL algebroids discussed in Lemma 3.12.

Control systems on AL algebroids

Definition 4.1. A *control system* on an AL algebroid E is a continuous map

$$f : M \times U \longrightarrow E \tag{4.1}$$

such that, for every $u \in U$, the map $f(\cdot, u) : M \longrightarrow E$ is a C^1 -section of E . We will assume that U is a subset of some Euclidean space \mathbb{R}^r . Moreover, we demand that the maps $f : M \times U \longrightarrow E$ and $T_x f : TM \times U \longrightarrow TE$ are continuous. In local coordinates, if $f \sim (f^i(x, u), x^a)$, this means that $f^i(x, u)$ is continuous w.r.t. x and u , differentiable w.r.t. x , and that $\frac{\partial f^i}{\partial x^a}(x, u)$ is continuous w.r.t. x and u .

Observe that, for the tangent algebroid $E = TM \rightarrow M$, the above definition coincides with the classical one (cf. Definition C.1 in Appendix C). On the other hand, one easily sees (cf. Theorem 3.6) that a right-invariant control system on a Lie groupoid \mathcal{G} reduces to a system of the above form on the associated Lie algebroid $\mathcal{A}(\mathcal{G})$. For example, a right-invariant control system on a gauge groupoid $\mathcal{G}_P = P \times P/G$ of a principal bundle $G \rightarrow P \rightarrow M$ is determined by its values on a single leaf of \mathcal{G}_P^α canonically isomorphic to P . Consequently, it is equivalent to a G -invariant control system on P and reduces to a control system of the form (4.1) on the Atiyah algebroid $E = TP/G$. In particular, for a right-invariant system on a Lie group G , Definition 4.1 coincides with the reduced control system on its Lie algebra \mathfrak{g} as described in [Jurdjevic, 1997, Ch. 12].

Now, for a given function $u : I \longrightarrow U$ (*control*), the map (4.1) defines a first-order ODE on M ,

$$\dot{x}(t) = \rho(f(x(t), u(t))). \quad (4.2)$$

We will restrict our attention only to functions u of a certain class (called *admissible controls*). In this paper these are controls which are bounded and measurable, but one can think of smaller classes: piecewise continuous or piecewise constant functions. The set of all admissible controls will be denoted by \mathcal{U}_{adm} .

Clearly, if $u(\cdot)$ is admissible, the map $g(x, t) = \rho(f(x, u(t)))$ is differentiable w.r.t. x and measurable w.r.t. t , so the assumptions of Theorem B.12 hold. Consequently, we have the results of local existence and uniqueness for the solutions of (4.2). Observe that if $x(\cdot)$ is a solution of (4.2) for $u(\cdot) \in \mathcal{U}_{adm}$, then the path $f(x(\cdot), u(\cdot)) : I \longrightarrow E$ is a measurable E -path over $x(\cdot)$. This path will be called a *trajectory* of (4.2), whereas for the pair $(x(\cdot), u(\cdot))$ we will use the term *controlled pair*.

Observe that if $E = TM$ is a tangent algebroid, then the trajectory of the system (4.2) is the tangent lift of the trajectory of a corresponding system (C.1) on M . The notions of controlled pairs coincide in both cases.

Optimal control problems on AL algebroids

We introduce now a *cost function* $L : M \times U \longrightarrow \mathbb{R}$. We will assume the same regularity conditions for L as in the case of f , namely, that L is a continuous function on $M \times U$, which is of class C^1 w.r.t. the first variable and that the derivative $d_x L : TM \times U \longrightarrow \mathbb{R}$ is continuous. If now $(x(t), u(t))$, with $t \in I$, is a controlled pair for (4.2), we define the *total cost* of this pair to be $\int_{t_0}^{t_1} L(x(t), u(t)) dt$. Note that, since L is continuous, $u(t)$ is bounded measurable, and the interval $[t_0, t_1]$ is compact, the above integral is finite whenever the solution $x(t)$ exists. Now we can define optimal control problems for the data introduced above. These definitions may seem unnatural at first sight, yet we will motivate them in the next subsection.

Definition 4.2. For a control system (4.1) and a cost function L we define an *optimal control problem* (OCP) as follows:

$$\begin{aligned} &\text{minimise } \int_{t_0}^{t_1} L(x(t), u(t)) dt \text{ over all controlled pairs } (x, u) \text{ of (4.2) such that} \\ &\quad \text{the } E\text{-homotopy class of the trajectory } f(x(t), u(t)) \text{ equals } [\sigma], \end{aligned} \quad (\text{P})$$

where σ is a fixed E -path. The interval $[t_0, t_1]$ is to be determined as well.

Given two smooth algebroid morphisms $\Phi_0 : TS_0 \rightarrow E$ and $\Phi_1 : TS_1 \rightarrow E$, we can also define an *optimal control problem relative to* (Φ_0, Φ_1) as follows:

$$\begin{aligned} &\text{minimise } \int_{t_0}^{t_1} L(x(t), u(t)) dt \text{ over all controlled pairs } (x, u) \text{ of (4.2) such that} \\ &\quad \text{the relative } E\text{-homotopy class of } f(x(t), u(t)) \text{ equals } [\sigma] \bmod(\Phi_0, \Phi_1), \end{aligned} \quad (\text{P rel})$$

where σ and $[t_0, t_1]$ are as above.

Interpretation of the algebroid OCPs

Now we shall relate the OCPs (P) and (P rel) to the standard OCPs considered in control theory. Briefly speaking, the E -homotopy restrictions in the OCPs on AL algebroids play a role of boundary condition in standard problems.

Let us concentrate first on (P). Recall from Chapter 3 that, for an integrable algebroid $\mathcal{A}(\mathcal{G})$, we have interpreted $\mathcal{A}(\mathcal{G})$ -homotopy classes as the standard homotopy classes reduced from a single α -fibre \mathcal{G}_{y_0} of an integrating groupoid \mathcal{G} to the associated algebroid $\mathcal{A}(\mathcal{G})$. Moreover, at the beginning of this chapter we have interpreted control system (4.1) on an integrable algebroid $E = \mathcal{A}(\mathcal{G})$ as a reduction of a right-invariant control system on the groupoid \mathcal{G} (or on a single α -fibre \mathcal{G}_{y_0}). Consider now a control system on a manifold N ,

$$\mathcal{F} : N \times U \longrightarrow TN, \quad (4.3)$$

with the cost function $\mathcal{L} : N \times U \longrightarrow \mathbb{R}$, and let us compare the following two OCPs on N :

$$\begin{aligned} \text{minimise } \int_{t_0}^{t_1} \mathcal{L}(x(t), u(t)) dt \text{ over all controlled pairs } (x, u) \\ \text{satisfying } x(t_0) = x_0, \ x(t_1) = x_1, \end{aligned} \quad (\text{P}_1)$$

and

$$\begin{aligned} \text{minimise } \int_{t_0}^{t_1} \mathcal{L}(x(t), u(t)) dt \text{ over all controlled pairs } (x, u) \\ \text{for which the homotopy class of } x(t) \text{ equals } [\sigma], \end{aligned} \quad (\text{P}_2)$$

where σ is a fixed path in N joining x_0 and x_1 and the time interval $[t_0, t_1]$ is not determined.

Problem (P₁) is a standard OCP on the manifold N . On the other hand, (P₂) is equivalent to the OCP (P) on a tangent algebroid TN .

We may also think of N as of an α -fibre \mathcal{G}_{y_0} of a groupoid \mathcal{G} with the control system (4.3) and the cost \mathcal{L} being \mathcal{G} -equivariant, and such that they reduce to the control system (4.1) and the cost function $L : U \times M \longrightarrow \mathbb{R}$ on the associated algebroid $E = \mathcal{A}(\mathcal{G})$. Clearly, in this situation, problem (P) on $E = \mathcal{A}(\mathcal{G})$ is equivalent to problem (P₂) on N .

Now let us compare problems (P₁) and (P₂). First, note that every solution of (P₁) gives a solution of (P₂) for some $[\sigma] \in \Pi_1(N, x_0, x_1)$. On the other hand, if we know the solutions of (P₂) for all possible classes $[\sigma] \in \Pi_1(N, x_0, x_1)$ then one (or more) of these solutions which has a minimal total cost is a solution of (P₁). To sum up, problem (P₂) is more refined than (P₁).

Observe that candidates for the solutions of (P₁) are usually indicated by the PMP. In the proof one compares the optimal trajectory with nearby (and hence homotopic) ones. Consequently, the PMP gives only conditions for local optimality and as such will also indicate all candidates for the solutions of (P₂) for all possible classes $[\sigma] \in \Pi_1(N, x_0, x_1)$ (if such candidates exist). Then, to solve (P₁) or (P₂), one has to investigate closer these candidates to check whether they are really optimal.

Finally, note that problems (P₁) and (P₂) are equivalent if N is simply connected. In fact, we can always lift the control system $\mathcal{F} : N \times U \rightarrow TN$ and the cost function $\mathcal{L} : N \times U \rightarrow \mathbb{R}$ to $\tilde{\mathcal{F}} : \tilde{N} \times U \rightarrow T\tilde{N}$ and $\tilde{\mathcal{L}} : \tilde{N} \times U \rightarrow \mathbb{R}$ defined on the universal cover \tilde{N} of N .

The discussion for the OCP (P rel) is quite similar. In the same setting as before consider two smooth maps $\tilde{\Phi}_0 : S_0 \rightarrow N$ and $\tilde{\Phi}_1 : S_1 \rightarrow N$. Now compare the following two OCPs on N :

$$\begin{aligned} & \text{minimise } \int_{t_0}^{t_1} \mathcal{L}(x(t), u(t)) dt \text{ over all controlled pairs } (x, u) \\ & \text{satisfying } x(t_0) \in \text{Im } \tilde{\Phi}_0, x(t_1) \in \text{Im } \tilde{\Phi}_1, \end{aligned} \quad (\text{P}_3)$$

and

$$\text{minimise } \int_{t_0}^{t_1} \mathcal{L}(x(t), u(t)) dt \text{ over all controlled pairs } (x, u) \text{ for which} \quad (\text{P}_4)$$

the homotopy class of $x(t)$ equals $[\sigma]$ relatively to the images $\text{Im } \tilde{\Phi}_0$ and $\text{Im } \tilde{\Phi}_1$.

Here σ is a fixed path in N , and we say that two paths σ_0 and σ_1 are homotopic relatively to the images $\text{Im } \tilde{\Phi}_0$ and $\text{Im } \tilde{\Phi}_1$ if there exists a homotopy joining σ_0 and σ_1 with the end-points in $\text{Im } \tilde{\Phi}_0$ and $\text{Im } \tilde{\Phi}_1$. As before the time interval $[t_0, t_1]$ is not fixed.

Problem (P₃) has a form of the standard OCP on the manifold N (one usually assumes that $\tilde{\Phi}_0$ and $\tilde{\Phi}_1$ are immersions). Problem (P₄), in turn, is equivalent to the OCP (P rel) for a control system on the tangent algebroid TN for algebroid morphisms $\Phi_0 = T\tilde{\Phi}_0 : TS_0 \rightarrow TN$ and $\Phi_1 = T\tilde{\Phi}_1 : TS_1 \rightarrow TN$.

Analogously as before, we can also think of N as of an α -fibre \mathcal{G}_{y_0} of a groupoid \mathcal{G} , with the control system (4.3) and the cost function being \mathcal{G} -equivariant and reducing to (4.2) and L . If now $\tilde{\Phi}_0$ and $\tilde{\Phi}_1$ are algebroid morphisms $\Phi_0 : TS_0 \rightarrow \mathcal{A}(\mathcal{G})$ and $\Phi_1 : TS_1 \rightarrow \mathcal{A}(\mathcal{G})$ lifted to $\mathcal{G}_{y_0} = N$ (cf. Corollary A.4), then (P₄) on N is equivalent to (P rel) on $E = \mathcal{A}(\mathcal{G})$.

Relation between (P₃) and (P₄) is analogous to the relation of (P₁) and (P₂):

- Every solution of (P₃) is a solution of (P₄) for some class $[\sigma] \bmod(\tilde{\Phi}_0, \tilde{\Phi}_1)$.
- The solution of (P₃) is this solution of (P₄) which has a minimal total cost of all solutions of (P₄) for all possible classes $[\sigma] \bmod(\tilde{\Phi}_0, \tilde{\Phi}_1)$.
- The solutions of (P₃) and (P₄) are not distinguishable by the PMP.

Problems (P₃) and (P₄) are equivalent if N is simply connected. This fact may not be obvious at first. It can be deduced from the following lemma.

Lemma 4.3. *Let N be a simply connected manifold, and let $S_0, S_1 \subset N$ be two path-connected subsets. Choose paths $\sigma_0, \sigma_1 : [0, 1] \rightarrow N$ such that $\sigma_0(0), \sigma_1(0) \in S_0$ and $\sigma_0(1), \sigma_1(1) \in S_1$. Then there exists a homotopy in N joining σ_0 with σ_1 which has its end-points in S_0 and S_1 .*

Proof. By path-connectedness of S_0 and S_1 , there exists a path $b_0 : [0, 1] \rightarrow S_0$ joining $\sigma_0(0)$ and $\sigma_1(0)$, and a path $b_1 : [0, 1] \rightarrow S_1$ joining $\sigma_0(1)$ and $\sigma_1(1)$. Denote by $\tilde{b}_0(t) := b_0(1 - t)$ and $\tilde{b}_1(t) := b_1(1 - t)$ the inverse paths of b_0 and b_1 .

Now path σ_0 and the composition $b_0 * \sigma_1 * \tilde{b}_1$ are homotopic with fixed-end-points in N , since they have the same end-points and N is simply connected (by $*$ we denote the concatenation of paths). Let $H : [0, 1] \times [0, 1] \rightarrow N$ be the appropriate homotopy. Consider homotopies $H_0 : [0, 1] \times [0, 1] \rightarrow S_0$ and $H_1 : [0, 1] \times [0, 1] \rightarrow S_1$ defined by the formulae $H_0(t, s) := b_0((1 - t)s)$ and $H_1(t, s) := b_1(ts)$. It is straightforward to verify that the composition of homotopies $\tilde{H} := H_0 * H * H_1$ makes sense, and it is a homotopy joining $\hat{\sigma}_0 := c_{\sigma_0(0)} * \sigma_0 * c_{\sigma_0(1)}$ with $\hat{\sigma}_1 := \tilde{b}_0 * b_0 * \sigma_1 * \tilde{b}_1 * b_1$ (here c_x stands for a constant path equal $x \in N$). Moreover, the initial-point homotopy of \tilde{H} is $H_0(0, s) = b_0(s) \in S_0$ and the final-point homotopy is $H_1(1, s) = b_1(s) \in S_1$, i.e., the paths $\hat{\sigma}_0$ and $\hat{\sigma}_1$ are homotopic relative to S_0 and S_1 .

To finish the proof observe that $\hat{\sigma}_0$ is homotopic (with fixed end-points) to σ_0 and $\hat{\sigma}_1$ to σ_1 . \square

OCPs in terms of the product algebroid $E \times \mathbb{T}\mathbb{R}$

For a control system on an AL algebroid E , similar to the classical situation of the tangent algebroid $E = TM \rightarrow M$, there is an elegant formulation of the OCPs (P) and (P rel) in terms of the product algebroid $E \times \mathbb{T}\mathbb{R}$. The idea is to incorporate the cost function into the control system (4.2).

Denote by \mathbf{A} the product algebroid structure on $\tau = (\tau_E, \tau_{\mathbb{T}\mathbb{R}}) : E \times \mathbb{T}\mathbb{R} \rightarrow M \times \mathbb{R}$ (see Chapter 2). We will consequently use bold letters to emphasise objects associated with \mathbf{A} , whereas objects associated with the $\mathbb{T}\mathbb{R}$ -component of \mathbf{A} will be distinguished by underlining. For example, $\mathbf{x} = (x, \underline{x}) \in M \times \mathbb{R}$ and $\mathbf{a} = (a, \underline{a}) \in E \times \mathbb{T}\mathbb{R} = \mathbf{A}$.

Introduce now a new variable $\underline{x} \in \mathbb{R}$ and, for a given admissible control $u \in \mathcal{U}_{adm}$, consider the following extension of the differential equation (4.2):

$$\begin{cases} \dot{x}(t) = \rho(f(x(t), u(t))), \\ \dot{\underline{x}}(t) = L(x(t), u(t)). \end{cases} \quad (4.4)$$

Clearly, $\underline{x}(t_1) - \underline{x}(t_0) = \int_{t_0}^{t_1} L(x(t), u(t)) dt$ is the total cost of the controlled pair $(x(t), u(t))$ of (4.2). Equation (4.4) is a differential equation associated with the following control system on \mathbf{A} :

$$\mathbf{f} = (\tilde{f}, \underline{f}) : (M \times \mathbb{R}) \times U \rightarrow E \times \mathbb{T}\mathbb{R} = \mathbf{A}, \quad (4.5)$$

where $\underline{f}((x, \underline{x}), u) := (\underline{x}, L(x, u)) \in \mathbb{R} \times \mathbb{R} \approx \mathbb{T}\mathbb{R}$ and $\tilde{f}((x, \underline{x}), u) := f(x, u) \in E$. For a given $u \in \mathcal{U}_{adm}$ the base trajectory $\mathbf{x}(t) = (x(t), \underline{x}(t))$ of (4.4) contains information on both the base trajectory $x(t)$ of (4.2) (for the same control u) and the total cost of the controlled pair $(x(t), u(t))$. Observe that the trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$ of (4.4) projects onto the trajectory $f(x(t), u(t))$ of (4.2) under the canonical algebroid projection $p_E : \mathbf{A} = E \times \mathbb{T}\mathbb{R} \rightarrow E$. Now the OCP (P) can be reformulated in terms of control system (4.5)

as follows:

- minimise $\underline{x}(t_1)$ over controlled pairs $(\mathbf{x}(t), u(t)) = ((x(t), \underline{x}(t)), u(t))$
of (4.4) satisfying the following:
1. the E -projection $f(x(t), u(t))$ of the trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$ (P)
belongs to a fixed E -homotopy class $[\sigma]$;
 2. $\underline{x}(t_0) = 0$.

Similarly, the OCP (P rel) can be expressed in the following way:

- minimise $\underline{x}(t_1)$ over controlled pairs $(\mathbf{x}(t), u(t)) = ((x(t), \underline{x}(t)), u(t))$
of (4.4) satisfying the following:
1. the E -projection $f(x(t), u(t))$ of the trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$ (P rel)
belongs to a fixed relative E -homotopy class $[\sigma] \bmod(\Phi_0, \Phi_1)$;
 2. $\underline{x}(t_0) = 0$.

The advantage of these new formulations of the OCPs (P) and (P rel) presented here may seem unclear. The main reason is that the unified treatment of cost and controls simplifies some aspects of the proof of the PMP.

The algebroid homotopy associated with a control system

As has been observed in [Crainic & Fernandes, 2003], algebroid homotopies can be generated by time-dependent algebroid sections. Since the control system (4.2) is a family of E -sections $f(\cdot, u)$, fixing an admissible control $u(t)$ gives a time-dependent section $f(\cdot, u(t))$. The associated E -homotopy can be well understood in terms of Lemma 3.12.

Solving (4.2) for a one-parameter family of initial conditions $x(t_0, s) = x_0(s)$ produces a one-parameter family of base paths $x(t, s)$. It follows from Theorem B.12 that, if the solution $x(t, 0)$ is defined on $I = [t_0, t_1]$, then so is $x(t, s)$ at least for $x_0(s)$'s close enough to $x_0(0)$. With $x(t, s)$ we can associate a one-parameter family of trajectories

$$a(t, s) := f(x(t, s), u(t)).$$

One easily sees that for (4.2) the assumptions of Theorem B.14 are satisfied. Consequently, the base trajectories $x(t, x_0)$ are continuous differentiable w.r.t. the initial condition x_0 (and ACB in t). As $x(t, s) = x(t, x_0(s))$ if $s \mapsto x_0(s)$ is an ACB map, we deduce that $x(t, s)$ is ACB w.r.t. the second variable; that is, $\partial_s x(t, s)$ is a well-defined measurable function of both variables. Consequently, the derivative $\partial_s a^i(t, s) = \frac{\partial f^i}{\partial x^a}(x(t, s), u(t)) \partial_s x^a(t, s)$ satisfies the assumptions of Lemma 3.12. Thus, the conclusions of Lemma 3.12 hold; namely, for a given bounded measurable E -path $b_0(s)$ covering $x_0(s)$, there exists a measurable map (AC w.r.t. the first variable) $b : I \times [0, 1] \rightarrow E$ with $b(t_0, s) = b_0(s)$ such that (a, b) is an E -homotopy. The t -evolution of $b(t, s)$ is given by (3.6). Observe that, since $\partial_s x(t, s) = \rho(b(t, s))$, we have $\partial_s a^i(t, s) = \frac{\partial f^i}{\partial x^a}(x(t, s), u(t)) \rho_k^a(x(t, s)) b^k(t, s)$. Consequently, $b(t, s) \sim$

$(x^a(t, s), b^i(t, s))$ is a solution of the following differential equation

$$\begin{cases} \partial_t b^i(t, s) = \frac{\partial f^i}{\partial x^a}(x(t, s), u(t)) \rho_k^a(x(t, s)) b^k(t, s) \\ \quad + c_{jk}^i(x(t, s)) b^j(t, s) f^k(x(t, s), u(t)), \\ \partial_t x^a(t, s) = \rho_i^a(x(t, s)) f^i(x(t, s), u(t)), \end{cases} \quad (4.6)$$

with the initial conditions $b^i(t_0, s) = b_0^i(s)$ and $x^a(t_0, s) = x_0^a(s)$.

The above differential equation is well understood in terms of the tools introduced in Chapter 2. For every $u \in U$, the section $f_u(\cdot) := f(\cdot, u) : M \rightarrow E$ gives rise to a linear vector field $d_T(f_u)$ on E . Evaluating it on $u(t)$ gives a time-dependent family of vector fields $d_T(f_{u(t)})$. Equation (4.6) is simply the evolution along this family, $\partial_t b(t, s) = d_T(f_{u(t)})(b(t, s))$. On the other hand, with a time-dependent family of section $f_u(t)$ we may associate the family of linear functions $h_t(x, \xi) := \langle f(x, u(t)), \xi \rangle_\tau$ on E^* , and the corresponding family of Hamiltonian vector fields \mathcal{X}_{h_t} . In local coordinates,

$$\mathcal{X}_{h_t}(x, \xi) = \rho_j^b(x) f^j(x, u(t)) \partial_{x^b} + \left(c_{ij}^k(x) f^i(x, u(t)) \xi_k - \rho_j^a(x) \frac{\partial f^i}{\partial x^a}(x, u(t)) \xi_i \right) \partial_{\xi_j}.$$

As we have seen in Chapter 2 (equations (2.4)–(2.6)), the fields $d_T(f_{u(t)})$ and \mathcal{X}_{h_t} give the same base evolution (given by (4.2)), and are related by $\langle d_T(f_{u(t)}), \mathcal{X}_{h_t} \rangle_{T\tau} = 0$.

Definition 4.4. The flows of the fields $d_T(f_{u(t)})$ and \mathcal{X}_{h_t} (for a given $u \in \mathcal{U}_{adm}$) will be called *operators of parallel transport* (in E and E^* respectively) along the solution $x(t)$ of the system (4.1). We will denote them with B_{tt_0} and $B_{tt_0}^*$, respectively. Analogously we define operators \mathbf{B}_{tt_0} and $\mathbf{B}_{tt_0}^*$ for the control system (4.5). Note that, by construction, $B_{tt'} \circ B_{t't_0} = B_{tt_0}$ and $B_{tt'}^* \circ B_{t't_0}^* = B_{tt_0}^*$.

Remark 4.5. Let us see that, by construction, the map $b(t, s) = B_{tt_0}(b_0(s))$ together with $a(t, s)$ forms an E -homotopy. Moreover, $B_{tt_0}(b_0)$ is continuous w.r.t. b_0 , t , and t_0 . Indeed, $b(t) = B_{tt_0}(b_0)$ is the solution of (4.6) for $s = 0$. The right-hand side is measurable in t and locally Lipschitz (linear) in b , so, by Theorem B.13, $b(t)$ is AC w.r.t. t and continuous w.r.t. the initial condition b_0 .

Remark 4.6. Note also that the operators B and B^* have the property of preserving the pairing $\langle \cdot, \cdot \rangle_\tau$; that is, for every $a \in E_{x(t')}$ and $\xi \in E_{x(t')}^*$ over the same base point $x(t') \in M$,

$$\langle B_{tt'}(a), B_{tt'}^*(\xi) \rangle_\tau = \langle a, \xi \rangle_\tau \quad \text{for every } t \in I.$$

Indeed, since by definition the pairing $\langle \cdot, \cdot \rangle_{T\tau} : TE \times_{TM} TE^* \rightarrow \mathbb{R}$ is the tangent map of $\langle \cdot, \cdot \rangle_\tau : E \times_M E^* \rightarrow \mathbb{R}$, we have

$$\begin{aligned} \partial_t \langle B_{tt'}(a), B_{tt'}^*(\xi) \rangle_\tau &= \langle \partial_t B_{tt'}(a), \partial_t B_{tt'}^*(\xi) \rangle_{T\tau} \\ &= \langle d_T(f_{u(t)})(B_{tt'}(a)), \mathcal{X}_{h_t}(B_{tt'}^*(\xi)) \rangle_{T\tau} = 0. \end{aligned}$$

Finally, observe that the evolution of $\boldsymbol{\xi}(t) = \mathbf{B}_{tt_0}^*(\boldsymbol{\xi}_0)$ for the control system (4.5) is trivial on the $T^*\mathbb{R}$ -component. Indeed, the associated linear Hamiltonian

$$\mathbf{H}_t(\mathbf{x}, \boldsymbol{\xi}) = \langle \mathbf{f}(\mathbf{x}, u(t)), \boldsymbol{\xi} \rangle_\tau = \langle f(x, u(t)), \xi \rangle_\tau + \underline{\xi} L(x, u(t))$$

does not depend on the \mathbb{R} -component of $\mathbf{x} = (x, \underline{x}) \in M \times \mathbb{R}$; hence $\boldsymbol{\xi}(t) = (\xi(t), \underline{\xi}(t)) \in \mathbf{A}^* = E^* \times T^*\mathbb{R}$ (in local coordinates, $(\xi, \underline{\xi}) \sim ((x^a, \xi_i), (\underline{x}, \underline{\xi}))$) evolves due to equations

$$\left\{ \begin{array}{l} \partial_t \xi_k(t) = -\rho_k^a(x) \left(\frac{\partial f^i}{\partial x^a}(x, u(t)) \xi_i(t) + \frac{\partial L}{\partial x^a}(x, u(t)) \underline{\xi}(t) \right) \\ \quad + c_{jk}^i(x) f^j(x, u(t)) \xi_i(t), \\ \partial_t \underline{\xi}(t) = 0, \\ \dot{x}(t) = \rho(f(x(t), u(t))), \\ \dot{\underline{x}}(t) = L(x(t), u(t)). \end{array} \right. \quad (4.7)$$

In other words, $\underline{\xi}(t) \equiv \underline{\xi}_0$ is a constant and $\xi(t) \in E^*$ evolves due to a time-dependent family of Hamiltonian vector fields \mathcal{X}_{H_t} on E^* , where $H_t(x, \xi) = \langle f(x, u(t)), \xi \rangle_\tau + \underline{\xi}_0 L(x, u(t))$.

The Pontryagin Maximum Principle

In the previous chapter we have introduced OCPs (P) and (P rel) in the AL algebroid setting. The main difference in comparison with the classical formulation are, apart from using algebroid-valued velocities, the fixed-homotopy boundary conditions. This new formulation of the OCPs was motivated in the previous chapter for problems on an integrable algebroid $E = \mathcal{A}(\mathcal{G})$. In light of these considerations we may think of (P) and (P rel) as of a general framework which includes the standard OCPs as well as OCPs reduced by general (groupoid) symmetries. It contains also abstract problems on general AL algebroids. Now we will formulate a generalisation of the Pontryagin maximum principle for the problems (P) and (P rel).

Formulation on an AL algebroid E

Theorem 5.1. *Let $(x(t), u(t))$, with $t \in [t_0, t_1]$, be a controlled pair of (4.2) solving the optimal control problem (P). Then there exists a curve $\xi : [t_0, t_1] \rightarrow E^*$ covering $x(t)$ and a constant $\underline{\xi}_0 \leq 0$ such that the following holds:*

- *the curve $\xi(t)$ is a trajectory of the time-dependent family of Hamiltonian vector fields \mathcal{X}_{H_t} associated with Hamiltonians $H_t(x, \xi) := H(x, \xi, u(t))$, where*

$$H(x, \xi, u) = \langle f(x, u), \xi \rangle_\tau + \underline{\xi}_0 L(x, u);$$

- *the control u satisfies the “maximum principle”*

$$H(x(t), \xi(t), u(t)) = \sup_{v \in U} H(x(t), \xi(t), v)$$

and $H(x(t), \xi(t), u(t)) = 0$ at every regular point t of u ;

- *if $\underline{\xi}_0 = 0$, the covector $\xi(t)$ is nowhere-vanishing.*

The above result clearly reduces to the standard PMP (Theorem C.2) for the case of the tangent algebroid $E = TM \rightarrow M$. It also covers the known results for system with symmetry on Lie groups and, more generally, principal bundles. A more detailed discussion and examples will be given in Chapter 6.

Consider smooth algebroid morphisms $\Phi_0 : TS_0 \rightarrow E$ over $\phi_0 : S_0 \rightarrow M$ and $\Phi_1 : TS_1 \rightarrow E$ over $\phi_1 : S_1 \rightarrow M$. For a relative OCP (P rel) we have the following version of the PMP.

Theorem 5.2. *Let $(x(t), u(t))$, with $t \in [t_0, t_1]$, $x(t_0) = \phi_0(z_0)$, and $x(t_1) = \phi_1(w_0)$, be a controlled pair of (4.2) solving the optimal control problem (P rel). Then there exists a curve $\xi : [t_0, t_1] \rightarrow E^*$ covering $x(t)$ and a constant $\underline{\xi}_0 \leq 0$ which satisfy the assertion of Theorem 5.1 and, additionally $\xi(t_0)$ annihilates $\Phi_0(T_{z_0}S_0)$ and $\xi(t_1)$ annihilates $\Phi_1(T_{w_0}S_1)$.*

It is clear that Theorem 5.2 agrees with the standard PMP for problems with general boundary conditions (Theorem C.3) for the special case of the tangent algebroid $E = TM$. We can regard this result as an extension of the standard PMP to systems with general (groupoid) symmetries. In particular, it covers the known results for symmetric systems on Lie groups and principal bundles (see Section 6.2).

Obviously, Theorem 5.1 is a special case of Theorem 5.2 obtained for Φ_0 and Φ_1 trivial.

Alternative formulation

Theorems 5.1 and 5.2 have equivalent formulations in terms of the product algebroid $A = E \times T\mathbb{R}$.

Theorem 5.3. *Let $(\mathbf{x}(t), u(t))$, with $t \in [t_0, t_1]$, be a controlled pair of (4.4) solving the optimal control problem (P). There exists a nowhere-vanishing curve $\underline{\xi} = (\xi, \underline{\xi}) : [t_0, t_1] \rightarrow A^* = E^* \times T^*\mathbb{R}$ covering $\mathbf{x}(t)$, with $\underline{\xi}(t_1) \leq 0$, such that the following hold:*

- *the curve $\underline{\xi}(t)$ is a trajectory of the time-dependent family of Hamiltonian vector fields \mathcal{X}_{H_t} , for $H_t(\mathbf{x}, \underline{\xi}) := H(\mathbf{x}, \xi, u(t))$, where*

$$H(\mathbf{x}, \xi, u) = \langle f(\mathbf{x}, u), \xi \rangle_\tau;$$

- *the control u satisfies the “maximum principle”*

$$H(\mathbf{x}(t), \underline{\xi}(t), u(t)) = \sup_{v \in U} H(\mathbf{x}(t), \underline{\xi}(t), v) = 0$$

at every regular point t of u .

The equivalence of Theorems 5.1 and 5.3 is obvious in light of our previous considerations. Indeed, the covector $\underline{\xi}(t)$ can be decomposed as $(\xi(t), \underline{\xi}(t))$, and its evolution along \mathcal{X}_{H_t} is given by (4.7). Consequently, as we have observed at the very end of Chapter 4, covector $\underline{\xi}(t) = \underline{\xi}_0$ is constant and the evolution of $\xi(t)$ is given by \mathcal{X}_{H_t} . Since $H(x, \xi, u) + \underline{\xi}_0 L(x, u) = H(\mathbf{x}, \underline{\xi}, u)$ for $\underline{\xi} = (\xi, \underline{\xi}_0)$, and $\mathbf{x} = (x, \underline{x})$, the corresponding statements in Theorems 5.1 and 5.3 are equivalent.

Define now $S_0 := \Phi_0(T_{z_0}S_0) \subset E_{x(t_0)}$ and $S_1 := \Phi_1(T_{w_0}S_1) \subset E_{x(t_1)}$, where $x(t_0) = \phi_0(z_0)$ and $x(t_1) = \phi_1(w_0)$. We have the following reformulation of Theorem 5.2.

Theorem 5.4. *Let $(\mathbf{x}(t), u(t))$, with $t \in [t_0, t_1]$ be a controlled pair of (4.4) solving the optimal control problem (P rel). Then there exists a nowhere-vanishing curve $\underline{\xi} : [t_0, t_1] \rightarrow A^*$ covering $\mathbf{x}(t)$, which satisfies the assertion of Theorem 5.3 and, additionally, $\underline{\xi}(t_0)$ annihilates $S_0 \oplus \theta_{\underline{x}(t_0)}$ and $\underline{\xi}(t_1)$ annihilates $S_1 \oplus \theta_{\underline{x}(t_1)}$.*

The equivalence of Theorems 5.4 and 5.3 is clear.

Remark 5.5. There are many different versions of the PMP — for autonomous and non-autonomous systems, with moving or fixed end-points, with free or fixed time interval $[t_0, t_1]$, etc. Theorem 5.4, in fact, covers all these cases (under certain regularity conditions). For details see Section 6.1.

Discussion of the main result

This part is devoted to the discussion of our main results formulated in the previous chapter. We begin with formulating and proving a version of the PMP on AL algebroids for non-autonomous systems. In section 6.2 we formulate a version of the PMP on for invariant OCPs on Lie groups and principal bundles. We derive Montgomery's falling cat problem as an illustration. Later we use the PMP to derive the known results from the calculus of variation on Lagrangian reduction, Hammel equations and Euler–Poincaré equations. We also formulate an algebroid analog of Euler-Lagrange equations. Finally, in section 6.3 we give a few concrete examples of the usage of our results.

6.1 Non-autonomous versions of the PMP

In [Pontryagin *et al.*, 1962] analogs of the PMP for other versions of the classical OCP (\tilde{P}) (including fixing the time interval, or changing the setting to the non-autonomous systems) were obtained. These extensions were proved by a clever reformulation of a problem given in order to make it a special case of the already known solution. Now we perform similar derivations for the extensions of the problems (P) and (P rel).

Consider a non-autonomous version of the control system (4.2) on an AL algebroid E

$$\dot{x}(t) = \rho(f(x(t), t, u(t))), \quad (6.1)$$

where $f : M \times \mathbb{R} \times U \longrightarrow E$ is a time-dependent family of C^1 -sections of E . Moreover, we assume that f is continuous w.r.t. all variables, differentiable w.r.t. x and t , and that the derivative $T_{(x,t)}f$ is also continuous w.r.t. all variables.

Let $L : M \times \mathbb{R} \times U \longrightarrow \mathbb{R}$ be a non-autonomous *cost function* satisfying the same regularity assumptions as f . Per analogy to definitions introduced in Chapter 4 we will speak of *trajectories* $f(x(t), t, u(t))$ and *extended controlled pairs* $(x(t), t, x(t))$ of (6.1).

Consider now the following generalisation of the OCP (P) :

$$\begin{aligned} & \text{minimise } \int_{t_0}^{t_1} L(x(t), t, u(t)) dt \text{ over all extended controlled pairs } (x, t, u) \\ & \text{of (6.1) s.t. the } E\text{-homotopy class of the trajectory } f(x(t), t, u(t)) \text{ equals } [\sigma]. \end{aligned} \quad (P_{na})$$

Here $[\sigma]$ is a fixed E -homotopy class. We allow the time interval $[t_0, t_1]$ either to be fixed or to be unspecified.

We can define also a relative version of the above problem by substituting the fixed- E -homotopy class $[\sigma]$ by a fixed relative- E -homotopy class $[\sigma] \bmod(\Phi_0, \Phi_1)$, for a pair of smooth algebroid morphisms $\Phi_0 : TS_0 \longrightarrow E$ and $\Phi_1 : TS_1 \longrightarrow E$.

For the above non-autonomous OCPs we have the following extension of Theorem 5.1.

Theorem 6.1. *Let $(x(t), t, u(t))$, with $t \in [t_0, t_1]$, be an extended controlled pair of (6.1) solving the optimal control problem (P_{na}) . Then there exists a curve $\xi : [t_0, t_1] \longrightarrow E^*$ covering $x(t)$ and a constant $\xi_0 \leq 0$ such that the following holds:*

- *the curve $\xi(t)$ is a trajectory of the time-dependent family of Hamiltonian vector fields $\mathcal{X}_{H_t}^E$ associated with Hamiltonians $H_t(x, \xi) := H(x, t, \xi, u(t))$, where*

$$H(x, t, \xi, u) = \langle f(x, t, u), \xi \rangle_\tau + \xi_0 L(x, t, u);$$

- *the control u satisfies the “maximum principle”*

$$H(x(t), t, \xi(t), u(t)) = \sup_{v \in U} H(x(t), t, \xi(t), v)$$

and $H(x(t), t, \xi(t), u(t)) - \int_{t_0}^t \frac{\partial H}{\partial s}(x(s), s, \xi(s), u(s)) ds = \text{const}$ at every regular point t of u . If the time interval $[t_0, t_1]$ is unspecified then this constant is 0;

- *if $\xi_0 = 0$, the covector $\xi(t)$ is nowhere-vanishing.*

Proof. The proof requires a simple reformulation of the given data. Consider, namely, the control system

$$\tilde{f} = (f, 1) : M \times \mathbb{R} \times U \longrightarrow E \times TR$$

on the product algebroid $E \times TR$ with the associated base dynamics

$$\begin{aligned} \dot{x}(t) &= \rho(f(x(t), z(t), u(t))), \\ \dot{z}(t) &= 1. \end{aligned}$$

Here $(x, z) \in M \times \mathbb{R}$. The variable z plays a role of time. Indeed, regardless of the chosen admissible control $u(t)$, the solution of the equation $\dot{z}(t) = 1$ with the initial condition $z(t_0) = t_0$ gives $z(t) = t$.

Now, if the time interval $[t_0, t_1]$ is fixed, OCP (P_{na}) is equivalent to the OCP (P) for the control system \tilde{f} with unspecified time interval and the $E \times TR$ -homotopy class defined by $[\sigma]$ on E and $[t_0, t_1]$ on TR . Consequently, we can use Theorem 5.1 to obtain the necessary conditions for optimality.

For $(\xi, \alpha) \in E_x^* \times T_z^* \mathbb{R}$ we define the Hamiltonian

$$\begin{aligned} \tilde{H}(x, z, \xi, \alpha, u) &:= \langle \tilde{f}(x, z, u), (\xi, \alpha) \rangle + \xi_0 L(x, z, u) = \\ &= \langle f(x, z, u), \xi \rangle_\tau + \xi_0 L(x, z, u) + 1 \cdot \alpha =: H(\xi, z, u) + \alpha. \end{aligned}$$

Now the evolution of the Pontryagin covector $(\xi(t), \alpha(t))$ reads as

$$\begin{aligned}\dot{\xi}(t) &= \mathcal{X}_{H(\cdot, z, u(t))}^E(\xi(t)) \\ \dot{\alpha}(t) &= -\frac{\partial}{\partial z} H(\xi(t), z, u(t)).\end{aligned}$$

Since $z(t) = t$ and $\tilde{H}(x(t), z(t), \xi(t), \alpha(t), u(t)) = 0$ at regular t , we get

$$\alpha(t) = c - \int_{t_0}^t \frac{\partial}{\partial s} H(\xi(s), s, u(s)) ds$$

and

$$H(\xi(t), t, u(t)) + \alpha(t) = 0 \quad \text{at } t \text{ regular.}$$

The maximum principle for H follows directly from the maximum principle for \tilde{H} . Finally, if $\xi_0 = 0$ and $\xi(t) = 0$, we would have $H(\xi(t), t, u(t)) = 0$, and hence also $\alpha(t) = 0$ a.e., which is impossible. This proves the assertion.

The proof for the case of unspecified time-interval $[t_0, t_1]$ is analogous, yet instead of fixed-homotopy class boundary conditions we have to general boundary conditions associated with algebroid morphisms $\Phi_0 = (\theta_{x_0}, \text{id}) : \mathbb{T}\mathbb{R} \rightarrow E \times \mathbb{T}\mathbb{R}$ and $\Phi_1 = (\theta_{x_1}, \text{id}) : \mathbb{T}\mathbb{R} \rightarrow E \times \mathbb{T}\mathbb{R}$. The additional condition $c = 0$ now follows from the transversality conditions of Theorem 5.2 for $\alpha(t_0)$ and $\alpha(t_1)$. \square

6.2 The known results

The PMP on Lie groups and principal bundles

The already proven results on the Lie groupoid—Lie algebroid reduction of a control system and homotopy (cf. Theorem 3.6, Corollary 3.8, and Chapter 4) allow us to formulate the following result which can be understood as a general reduction scheme of the PMP on a principal G -bundle. Consider a principal G -bundle $G \rightarrow P \xrightarrow{\pi} M$.

Theorem 6.2. *Let $F : P \times U \rightarrow TP$ be a G -invariant control system on P and let $L : P \times U \rightarrow \mathbb{R}$ be a G -invariant cost function. Choose a path $\Sigma : [t_0, t_1] \rightarrow P$ joining two fixed points $p_0, p_1 \in P$.*

Then the OCP (P_2) on P for a fixed homotopy class $[\Sigma]$ in P is equivalent to the OCP (P) for the system $f : M \times U \rightarrow E = TP/G$ where $f(\pi(p), u) := [F(p, u)]$, with the cost function $l : M \times U \rightarrow \mathbb{R}$ defined by $l(\pi(p), u) := L(p, u)$, and the E -homotopy class $[\sigma]$ being the reduction of $[\Sigma]$.

Now applying Theorem 5.1 to the OCP described above we obtain a general result for equivariant OCP on principal bundles described in terms of the linear Poisson structure Π_{E^*} on $E^* = T^*P/G$ (cf. Theorem A.7).

Theorem 6.3. *Let $(p(t), u(t))$, with $t \in [t_0, t_1]$, be a controlled pair of F solving the OCP described above. Denote by $x(t)$ the base projection of $p(t)$. Then there exists a curve $\xi : [t_0, t_1] \rightarrow T^*P/G$ covering $x(t)$ and a constant $\underline{\xi}_0 \leq 0$ such that the following holds:*

- *the curve $\xi(t)$ is a trajectory of the time-dependent family of Hamiltonian vector fields \mathcal{X}_{h_t} associated with the linear Poisson structure Π_{E^*} on T^*P/G and Hamiltonians $h_t(x, \xi) := h(x, \xi, u(t))$, where*

$$h(x, \xi, u) = \langle f(x, u), \xi \rangle_\tau + \underline{\xi}_0 l(x, u);$$

- *the control u satisfies the “maximum principle”*

$$h(x(t), \xi(t), u(t)) = \sup_{v \in U} h(x(t), \xi(t), v)$$

and $h(x(t), \xi(t), u(t)) = 0$ at every regular point t of u ;

- *if $\underline{\xi}_0 = 0$, the covector $\xi(t)$ is nowhere-vanishing.*

An analogous result (with additional transversality conditions) is valid for system with general boundary conditions.

Note that for the case of a Lie group ($P = G$, $M = \{*\}$, $\Pi_{E^*} = \Pi_{\mathfrak{g}^*}$) we recover the results of Jurdjevic [Jurdjevic, 1997, Ch.12, Thms 5,6].

An application—the falling cat problem

Now we will reconsider the well-known results of Montgomery [Montgomery, 1990] (see also [Bloch, 2003, ch. 7.1] and [Cendra *et al.*, 1998]) on the isoholonomic problem by means of the PMP in the Atiyah algebroid setting.

Let $G \rightarrow P \rightarrow M$ be a principal G -bundle, let $\mathcal{H} \subset TP$ be a G -invariant horizontal distribution, and let $\mu(\cdot, \cdot)$ be a G -invariant sub-Riemannian metric on \mathcal{H} ($\mu(\cdot, \cdot)$ can be understood as a base metric lifted to \mathcal{H} by the horizontal lift). The problem is now to find a horizontal curve $q(t)$ with $t \in [0, 1]$ joining two fixed points q_0, q_1 in P and minimizing the total energy

$$\frac{1}{2} \int_0^1 \mu(\dot{q}(t), \dot{q}(t)) dt.$$

Clearly, due to the G -invariance of the problem, after changing the fixed-end-point condition into a fixed-homotopy condition (as discussed in detail in Chapter 4), the above problem is equivalent to an OCP of the form (P) on the Atiyah algebroid $E = TP/G$.

With the invariant distribution \mathcal{H} , understood as a principal connection, we can associate a map $\nabla : TM \rightarrow E$ inducing a splitting $E \simeq_{\nabla} TM \times \mathfrak{g}$. Our control system will be $f : TM \rightarrow TM \times \mathfrak{g}$ given by $f(X) = (X, 0)$ (this assures that the trajectory is horizontal), the cost function $L : TM \rightarrow \mathbb{R}$ reads as $L(X) = \frac{1}{2}\mu(X, X)$, and the fixed E -homotopy class is simply a reduction of a classical homotopy class in P .

Note two differences with the formulation of the OCP (P). Firstly, our control and cost functions have arguments in TM instead of in $M \times U$. Of course, this makes no big

difference, since locally $TM \approx M \times \mathbb{R}^n$. Secondly, our time interval is fixed. This, in turn, results in substituting the condition $H(x(t), \xi(t), u(t)) = 0$ by $H(x(t), \xi(t), u(t)) = \text{const}$ in the assertion of Theorem 5.1 (cf. Section 6.1).

Now we can apply Theorem 5.1 with the Hamiltonian evolution described in Theorem A.7. The covector $\xi \in E^*$ can be decomposed as $\xi = (p, \zeta) \in T^*M \times \mathfrak{g}^*$, and the corresponding Hamiltonian is

$$H(p, \zeta, X) = \langle X, p \rangle + \frac{1}{2} \lambda_0 \mu(X, X) =: h(p, X),$$

with $\lambda_0 \leq 0$. The maximum principle reads as $p(t) = -\lambda_0 \mu(X(t), \cdot)$; hence on the optimal trajectory, $H(p(t), \zeta(t), X(t)) = -\frac{1}{2} \lambda_0 \mu(X(t), X(t))$ (which is constant in t). The evolution of $p(t)$ and $\zeta(t)$ is given by

$$\begin{aligned} \dot{\zeta}(t) &= 0, \\ \dot{p}(t) &= \mathcal{X}_{h(p,X)}^{T^*M} + \langle \xi(t), F_{\nabla}(X, \cdot) \rangle; \end{aligned}$$

hence $\zeta(t) = \text{const}$. The second equation is equivalent to

$$\lambda_0 \nabla_X^\mu X = \langle \zeta, F_{\nabla}(X, \cdot) \rangle^{\# \mu},$$

where ∇^μ denotes the Levi-Civita covariant derivative on (M, μ) and $A^{\# \mu}$ is the vector dual to A by means of μ . Indeed, the equation $\dot{p}(t) = \mathcal{X}_{h(p,X)}^{T^*M}$, together with $p(t) = -\lambda_0 \mu(X(t), \cdot)$, is the PMP for a geodesic problem on (M, μ) . Passing to the dual vector $p(t)^{\# \mu} = \lambda_0 X(t)$ we should obtain the geodesic equation multiplied by the factor λ_0 . The equation $\dot{\zeta}(t) = 0$ means that the curve $\zeta(t) \in \mathfrak{g}^*$ is covariantly constant, hence

$$\nabla_X \zeta = 0.$$

We have thus obtained the Wong equations as in [Montgomery, 1990].

The abnormal case $\lambda_0 = 0$ implies $p(t) = 0$ and $\langle \zeta, F_{\nabla}(X, \cdot) \rangle = 0$. This allows us to exclude abnormal solutions in certain situations. For example, if P is a bundle of circles over a two-dimensional base and the connection is non-integrable (i.e., F_{∇} is non-vanishing), we have $\langle \zeta, F_{\nabla}(X, \cdot) \rangle = 0$ if and only if $X = 0$ (hence the solution is trivial) or $\zeta = 0$, which can be excluded by the non-vanishing of the covector in the PMP.

Applications to variational problems

It is a well-known fact that Euler-Lagrange equations can be derived by means of the classical PMP if one considers a trivial control system on a manifold M

$$f : M \times \mathbb{R}^n \underset{\text{loc}}{\approx} TM \xrightarrow{\text{id}} TM.$$

In this case, since we make no restrictions for velocities, the abnormal case can be excluded.

Similarly, for a trivial control system on a general AL algebroid

$$f : M \times \mathbb{R}^m \underset{\text{loc}}{\approx} E \xrightarrow{\text{id}} E,$$

we can obtain generalised Euler-Lagrange equations studied by many authors (see e.g. [Grabowska & Grabowski, 2008] and the references therein).

Indeed, for a system of the above form with a cost function

$$L : M \times \mathbb{R}^m \underset{\text{loc}}{\approx} E \longrightarrow \mathbb{R}$$

consider the OCP (P) for some fixed homotopy class $[\sigma]$ and fixed time interval $[t_0, t_1]$ (cf. Section 6.1), and denote by $\gamma : [t_0, t_1] \longrightarrow E$ its solution (the controlled pair). In local coordinates $\gamma(t) \sim (x^a(t), y^i(t))$. The associated time-dependent Hamiltonian $H : \mathbb{R} \times E^* \longrightarrow \mathbb{R}$ reads as

$$H(t, \xi) = \langle \gamma(t), \xi \rangle_\tau - \xi_0 L(\gamma(t)).$$

The maximum principle

$$H(t, \xi(t)) = \sup_{e \in E_{\pi(\xi(t))}} \langle e, \xi(t) \rangle + \xi_0 L(e)$$

implies that if $\xi_0 = 0$, then also $\xi(t) = 0$, which is forbidden by the PMP. Consequently, we may assume that $\xi_0 = -1$. Now the maximum condition implies that the Pontryagin covector $\xi(t)$ is a vertical part of the derivative dL evaluated on $\gamma(t)$; i.e., $\xi(t) = T^*\pi(dL(\gamma(t))) \sim \left(x^a(t), \frac{\partial L}{\partial y^i}(x(t), y(t))\right)$.

The evolution equation reads as

$$\dot{\xi}(t) = \tilde{\Pi}_{E^*}(d_\xi H(t, \xi(t))), \quad (6.2)$$

where $\tilde{\Pi}_{E^*} : T^*E^* \longrightarrow TE^*$ is induced by the linear bi-vector field Π_{E^*} .

For our purposes it will be more convenient to describe the dynamics via the canonical double vector bundle isomorphism $\mathcal{R}^{-1} : T^*E^* \longrightarrow T^*E$ (see [Mackenzie, 2005, Sec. 11]), which in local coordinates reads as

$$\mathcal{R}^{-1} : (x^a, \xi_i, p_b, y^j) \mapsto (x^a, y^i, -p_b, \xi_j).$$

Since $d_\xi H(t, \xi) \sim (x^a(t), \xi^i(t), -\frac{\partial L}{\partial x^b}(x(t), y(t)), y^i(t))$, the image $\mathcal{R}^{-1}(d_\xi H(t, \xi(t)))$ is simply the derivative dL evaluated at $\gamma(t)$. Equation (6.2) can be thus expressed as

$$\frac{d}{dt} T^*\tau(\gamma(t)) = \varepsilon \circ dL(\gamma(t)),$$

where $\varepsilon := \tilde{\Pi}_{E^*} \circ \mathcal{R} : T^*E \longrightarrow TE^*$. This equation considered as an implicit differential equation for $\gamma(t)$ is precisely the generalised *Euler-Lagrange equations* considered in [Grabowska & Grabowski, 2008, Grabowska *et al.*, 2006]. In local coordinates it reads as

$$\begin{aligned} \frac{dx^a}{dt} &= \rho_k^a(x) y^k \\ \frac{d}{dt} \left(\frac{\partial L}{\partial y^j} \right) &= c_{ij}^k(x) \frac{\partial L}{\partial y^k} + \rho_j^a(x) \frac{\partial L}{\partial x^a}. \end{aligned} \quad (6.3)$$

In a special case if $E = \mathfrak{g}$ is a Lie algebra we recover the *Euler–Poincaré equations*

$$\frac{d}{dt} \left(\frac{\partial L}{\partial y} \right) = \text{ad}_y^* \left(\frac{\partial L}{\partial y} \right)$$

More generally, for the Atiyah algebroid $E = TP/G$, generalised Euler–Lagrange equations (6.3) take a form of *Hammel equations* (if we use local trivialisation defined by a local section—see (A.3)) and *reduced Euler–Lagrange equations* (in local trivialisation given by a principal connection—see (A.4)) The interested reader should confront [Cendra *et al.*, 1998, Sec. 5].

6.3 Other examples

Two-point time OCP on $\mathfrak{so}(3)$

Consider now a rigid body in \mathbb{R}^3 which can rotate with constant angular velocity along two fixed axes in the body. At every moment the position of the body is described by an element $q \in SO(3)$. The rotation axes can be represented by elements of the Lie algebra $l_+, l_- \in \mathfrak{so}(3)$. The rotation along the axis l_{\pm} is described by the equation

$$\partial_t q = ql_{\pm}.$$

It would be suitable to write $l_+ = a + b$ and $l_- = a - b$. The above equation can be regarded as a control system on the Lie group $SO(3)$ with the control function $F(q, u) = qf(u)$, where $f(u) = a + ub$ and the set of controls is simply $U = \{-1, 1\}$. We would like to find a control $u(t)$ which moves the body from a position $q_0 \in SO(3)$ to $q_1 \in SO(3)$ (or such that the trajectory belongs to a fixed homotopy class in $SO(3)$) in the shortest possible time.

It is obvious that the above OCP on the Lie group reduces to the OCP on the Lie algebra $\mathfrak{so}(3)$ with the control function f and the cost function $L \equiv 1$. Fix a basis (e_1, e_2, e_3) on $\mathfrak{so}(3)$, and denote by $c_{\beta\gamma}^{\alpha}$ the structure constants of the Lie algebra in this basis. Let $u(t)$, for $t \in [t_0, t_1]$, be a solution of the above OCP. It follows from theorem 5.1 that there exist a number $\lambda_0 \leq 0$ and a curve $\zeta(t) \in \mathfrak{so}(3)^*$ such that

$$H(\zeta(t), u(t)) = \langle \zeta(t), a + u(t)b \rangle + \lambda_0 = \max_{v=\pm 1} \langle \zeta(t), a + vb \rangle + \lambda_0.$$

This implies that $u(t) = \text{sgn}(\langle \zeta(t), b \rangle)$. Moreover, the evolution of $\zeta(t)$ is given by the equation

$$\partial_t \zeta_{\beta}(t) = c_{\alpha\beta}^{\gamma}(a^{\alpha} + u(t)b^{\alpha})\zeta_{\gamma}(t).$$

We have obtained the same equation as in ([Agrachev & Sachkov, 2004, Sec. 19.4]). We refer the reader to this book for the detailed discussion on solutions.

An application to a nonholonomic system

In [Grabowska *et al.*, 2006] and [Grabowska & Grabowski, 2008] a framework of geometric mechanics on general algebroids was presented. Roughly speaking, the structure of an algebroid on a bundle $\tau : E \rightarrow M$ allows one to develop Lagrangian formalism for a given Lagrangian function $L : E \rightarrow \mathbb{R}$. Moreover, if E is an AL algebroid, then the associated Euler-Lagrange equations have a variational interpretation: a curve $\gamma : [t_0, t_1] \rightarrow E$ satisfies the Euler-Lagrange equations if and only if it is an extremal of the action $\mathcal{J}(\gamma) := \int_{t_0}^{t_1} L(\gamma(t))dt$ restricted to those γ 's which are admissible and belong to a fixed E -homotopy class [Grabowska & Grabowski, 2008]. Hence, the trajectories of the Lagrange system should be derivable from our version of the PMP for the unconstrained control system on E with the cost function L .

In [Grabowski *et al.*, 2009] it has been shown that if $D \subset E$ is a subbundle and L is of mechanical type (that is, $L(a) = \frac{1}{2}\mu(a, a) - V(\tau(a))$, where μ is a metric on E and V is an arbitrary function on the base), then nonholonomically constrained Euler-Lagrange equations associated with D can be obtained as unconstrained Euler-Lagrange equations on the skew-algebroid $(D, \rho_E|_D, [\cdot, \cdot]_D := P_D[\cdot, \cdot]_E)$, where $P_D : E \rightarrow D$ denotes the projection orthogonal w.r.t. μ . It follows that if D with the algebroid structure defined above is AL, then the solutions of the nonholonomically constrained Euler-Lagrange equations are extremals of the unconstrained OCP on D with the cost function $L|_D$. On the other hand, using our version of the PMP on the algebroid E with controls restricted to D and the cost function L , one will obtain nonholonomically constrained Euler-Lagrange equations associated with D . Note that the algebroid bracket $[\cdot, \cdot]_D$ need not satisfy Jacobi identity even if $[\cdot, \cdot]_E$ does. Concluding, the PMP on general (not necessarily Lie) AL algebroids can be used in the theory of nonholonomic systems. To our knowledge this point of view is completely novel.

To give a concrete example we will use PMP to study the Chaplygin sleigh. It is an example of a nonholonomic system on the Lie algebra $\mathfrak{se}(2)$ which describes a rigid body sliding on a plane. The body is supported in three points, two of which slide freely without friction, while the third point is a knife edge. This imposes the constraint of no motion orthogonal to this edge (see [Chaplygin, 1911, Neimark & Fufaev, 1972]).

The configuration space before reduction is the Lie group $G = SE(2)$ of the Euclidean motions of the two-dimensional plane \mathbb{R}^2 . Elements of the Lie algebra $\mathfrak{se}(2)$ are of the form

$$\hat{\xi} = \begin{pmatrix} 0 & \omega & v_1 \\ -\omega & 0 & v_2 \\ 0 & 0 & 0 \end{pmatrix} = v_1 E_1 + v_2 E_2 + \omega E_3,$$

where $[E_3, E_1] = E_2$, $[E_2, E_3] = E_1$, and $[E_1, E_2] = 0$.

The system is described by the purely kinetic Lagrangian function $L : \mathfrak{se}(2) \rightarrow \mathbb{R}$, which reads as

$$L(v_1, v_2, \omega) = \frac{1}{2} [(J + m(a^2 + b^2))\omega^2 + mv_1^2 + mv_2^2 - 2bm\omega v_1 - 2am\omega v_2].$$

Here m and J denote the mass and the moment of inertia of the sleigh relative to the contact point, while (a, b) represents the position of the centre of mass w.r.t. the body

frame, determined by placing the origin at the contact point and the first coordinate axis in the direction of the knife axis. Additionally, the system is subjected to the nonholonomic constraint determined by the linear subspace

$$D = \{(v_1, v_2, \omega) \in \mathfrak{se}(2) \mid v_2 = 0\} \subset \mathfrak{se}(2).$$

Instead of $\{E_1, E_2, E_3\}$ we take another basis of $\mathfrak{se}(2)$:

$$e_1 = E_3, \quad e_2 = E_1, \quad e_3 = -maE_3 - mabE_1 + (J + ma^2)E_2,$$

adapted to the decomposition $D \oplus D^\perp$; $D = \text{span} \{e_1, e_2\}$ and $D^\perp = \text{span} \{e_3\}$. The induced skew-algebroid structure on D is given by

$$[e_1, e_2]_D = \frac{ma}{J + ma^2}e_1 + \frac{mab}{J + ma^2}e_2.$$

Therefore, the structural constants are $\mathcal{C}_{12}^1 = \frac{ma}{J + ma^2}$ and $\mathcal{C}_{12}^2 = \frac{mab}{J + ma^2}$. The algebroid D is almost Lie (in fact, in this simple case it is a Lie algebra). Next, we will use theorem 5.1 to derive the nonholonomic equations of motion. Set $U = \mathbb{R}^2 \ni (y^1, y^2)$ and the control function to be a map $f : U \rightarrow D$ given by

$$f(y^1, y^2) = y^1 e_1 + y^2 e_2 \in D.$$

The Lagrangian restricted to D defines the cost function $L : U \rightarrow \mathbb{R}$,

$$L(y^1, y^2) = \frac{1}{2} [(J + m(a^2 + b^2))(y^1)^2 + m(y^2)^2 - 2bmy^1 y^2].$$

For a curve $\xi(t) = \xi_1(t)e_1^* + \xi_2(t)e_2^* \in D^*$ the maximum principle reads

$$\begin{aligned} H(\xi(t), y(t)) &= \xi_1(t)y^1 + \xi_2(t)y^2 + \underline{\xi}_0 \cdot L(y^1, y^2) \\ &= \max_{(v^1, v^2) \in \mathbb{R}^2} (\xi_1(t)v^1 + \xi_2(t)v^2 + \underline{\xi}_0 \cdot L(v^1, v^2)). \end{aligned} \quad (6.4)$$

If $\underline{\xi}_0 = 0$, then maximality would give $\xi(t) = 0$, which is impossible. Hence, we may assume that $\underline{\xi}_0 = -1$. Now from (6.4) we will get

$$\begin{aligned} \xi_1(t) &= (J + m(a^2 + b^2))y^1 - bmy^2, \\ \xi_2(t) &= my^2 - bmy^2. \end{aligned} \quad (6.5)$$

Finally, the Hamiltonian evolution (4.7) is simply

$$\begin{aligned} \dot{\xi}_1 &= \mathcal{C}_{21}^1 y^2 \xi_1 + \mathcal{C}_{21}^2 y^2 \xi_2 = -\frac{ma}{J + ma^2} y^2 (\xi_1 + b\xi_2), \\ \dot{\xi}_2 &= \mathcal{C}_{21}^1 y^1 \xi_1 + \mathcal{C}_{12}^2 y^1 \xi_2 = \frac{ma}{J + ma^2} y^1 (\xi_1 + b\xi_2). \end{aligned}$$

In view of (6.5) and the above equations we conclude that the equations of motion are

$$\begin{aligned} (J + m(a^2 + b^2))\dot{y}^1 - bmy^2 &= -may^1 y^2, \\ my^2 - bmy^1 &= ma(y^1)^2, \end{aligned}$$

which completely agrees with [Grabowski *et al.*, 2009].

Needle variations

In order to prove Theorems 5.3 and 5.4 we shall somehow compare the cost on the optimal trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$ with costs of nearby trajectories. As our assumptions input on the set of controls U are very mild, we cannot use the natural concept of a continuous deformation, as in the standard calculus of variations (U can be for instance discrete). Instead, we introduce the notion of *needle variations* after [Pontryagin *et al.*, 1962]. For a given admissible control $u : [t_0, t_1] \rightarrow U$ this variation will be, roughly speaking, the family of controls $u_s(t)$ obtained by substituting $u(t)$ by given elements $v_i \in U$ on a small intervals $I_i = (\tau_i - s\delta t_i, \tau_i] \subset [t_0, t_1]$. Our main result in this chapter is Theorem 7.4, where we study the \mathbf{A} -homotopy classes of the trajectories of the system (4.4) obtained for controls $u_s(t)$. We finish this chapter with the definition of \mathbf{K}_τ^u —the set of infinitesimal variations of the trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$.

Needle variation of controls and trajectories

Throughout this chapter we will work with a fixed admissible control $u : [t_0, t_1] \rightarrow U$ and fixed trajectory $\mathbf{a}(t) := \mathbf{f}(\mathbf{x}(t), u(t))$.

Choose points $t_0 < \tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq \tau < t_1$, being regular points of u . Next, choose non-negative numbers $\delta t_1, \dots, \delta t_k$ and an arbitrary real number δt . Finally, take (not necessarily different) elements $v_1, \dots, v_k \in U$. The whole set of data $(\tau_i, v_i, \delta t_i, \delta t)_{i=1, \dots, k}$ will be denoted by \mathbf{w} and called a *symbol*. Its role will be to encode the variation of the control $u(t)$. Intuitively, points τ_i emphasise moments in which we substitute $u(t)$ by a constant control v_i on an interval $I_i = (\tau_i - s\delta t_i, \tau_i]$ of length $s\delta t_i$, while $s\delta t$ is responsible for shortening or lengthening the time for which $u(t)$ is defined. The precise definition is quite technical, because one should take care to make the intervals I_i pair-wise disjoint.

Take

$$l_i = \begin{cases} \delta t - (\delta t_i + \dots + \delta t_k) & \text{when } \tau_i = \tau; \\ -(\delta t_i + \dots + \delta t_k) & \text{when } \tau_i = \tau_k < \tau; \\ -(\delta t_i + \dots + \delta t_j) & \text{when } \tau_i = \tau_{i+1} = \dots = \tau_j < \tau_{j+1}, \end{cases}$$

and define s -dependent intervals $I_i := (\tau_i + sl_i, \tau_i + s(l_i + \delta t_i)]$. As we see, I_i is left-open and right-closed and it has length $s \cdot \delta t_i$. If $\tau_i < \tau_{i+1}$, or $i = k$ and $\tau_k < \tau$, the end-point of I_i lies at τ_i . If $\tau_i = \tau_{i+1}$, then the end-point of I_i coincides with the initial-point of I_{i+1} . If $\tau_k = \tau$, we set the end-point of I_k at $\tau + s\delta t$. Clearly, for s small enough, the intervals I_i lie inside $[t_0, t_1]$ and are pairwise disjoint.

Definition 7.1. For a symbol $\mathfrak{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$ we introduce a s -dependent family of admissible controls defined on intervals $[t_0, t_1 + s\delta t]$:

$$u_s^{\mathfrak{w}}(t) = \begin{cases} v_i & \text{for } t \in I_i, \\ u(t) & \text{for } t \in [t_0, \tau + s\delta t] \setminus \bigcup_i I_i \\ u(t - s\delta t) & \text{for } t \in (\tau + s\delta t, t_1 + s\delta t]. \end{cases} \quad (7.1)$$

We will call $u_s^{\mathfrak{w}}$ a (*needle*) *variation of the control u associated with the symbol \mathfrak{w}* .

Using $u_s^{\mathfrak{w}}(t)$ and an AC path $s \mapsto \mathbf{x}_0(s) \in M \times \mathbb{R}$ where $\mathbf{x}_0(0) = \mathbf{x}_0$ we can define the variation of $\mathbf{a}(t)$.

Definition 7.2. The family of trajectories

$$\mathbf{a}(t, s) := \mathbf{f}(\mathbf{x}(t, s), u_s^{\mathfrak{w}}(t))$$

of the system (4.4), with the initial conditions $\mathbf{x}(t_0, s) = \mathbf{x}_0(s)$, where $t \in [t_0, t_1 + s\delta t]$, will be called a *variation of the trajectory $\mathbf{a}(t) = \mathbf{f}(\mathbf{x}(t), u(t))$ associated with the symbol \mathfrak{w} and the initial base-point variation $\mathbf{x}_0(s)$* .

Remark 7.3. Observe that, when $\delta t_i = 0$, the interval I_i is empty. It follows that adding a triple $(\tau_i, v_i, \delta t_i = 0)$ to the symbol \mathfrak{w} does not change the variation $u_s^{\mathfrak{w}}$ and, consequently, the associated variations $\mathbf{a}(t, s)$.

Needle variations and \mathbf{A} -homotopy classes

Our goal now is to compare the \mathbf{A} -homotopy classes of the trajectory $\mathbf{a}(t)$ and its variation $\mathbf{a}(t, s)$ introduced above. We need this because OCPs (P) and (P rel) are defined in term of algebroid homotopy classes. Having in mind Lemma 3.12 and the construction of a \mathbf{A} -homotopy associated with a control system (4.4) given in Chapter 4, we may expect that the family of trajectories $\mathbf{a}(t, s)$ forms an \mathbf{A} -homotopy (for some initial-point homotopy $\mathbf{b}_0(s)$). Consequently, the description of \mathbf{A} -homotopy classes of $\mathbf{a}(t, s)$ should be possible by means of Lemma 3.15. This is indeed the case, yet some technical work is needed in order to reparametrise $\mathbf{a}(t, s)$ in a suitable way.

Theorem 7.4. Let $s \mapsto \mathbf{b}_0(s)$ be a bounded measurable \mathbf{A} -path covering $s \mapsto \mathbf{x}_0(s)$, where $\mathbf{x}_0(0) = \mathbf{x}_0$. Consider a variation $\mathbf{a}(t, s) = \mathbf{f}(\mathbf{x}(t, s), u_s^{\mathfrak{w}}(t))$ of the trajectory $\mathbf{a}(t) = \mathbf{f}(\mathbf{x}(t), u(t))$ associated with a symbol $\mathfrak{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$ and initial base-point variation $\mathbf{x}_0(s)$.

Then there exists a number $\theta > 0$ and an \mathbf{A} -path $s \mapsto \mathbf{d}^{\mathfrak{w}}(s)$ defined for $0 \leq s \leq \theta$ such that

$$[\mathbf{b}_0(s)]_{s \in [0, \varepsilon]} [\mathbf{a}(t, \varepsilon)]_{t \in [t_0, t_1 + \varepsilon \delta t]} = [\mathbf{a}(t)]_{t \in [t_0, t_1]} [\mathbf{d}^{\mathfrak{w}}(s)]_{s \in [0, \varepsilon]}, \quad (7.2)$$

for every $\varepsilon \leq \theta$.

Moreover, if (τ_i, v_i, τ) in \mathfrak{w} are fixed, we can choose $\theta > 0$ universal for all $(\delta t_i, \delta t)$ belonging to a fixed compact set.

Finally, if $\mathbf{b}_0(s)$ is regular at $s = 0$, then $\mathbf{d}^{\mathfrak{w}}(s)$, regarded as a function of s , δt_i and δt , is uniformly regular w.r.t. δt_i and δt at $s = 0$. What is more,

$$\begin{aligned} \mathbf{d}^{\mathfrak{w}}(0) &= \mathbf{B}_{t_1\tau}[\mathbf{f}(\mathbf{x}(\tau), u(\tau))]\delta t + \mathbf{B}_{t_1t_0}(\mathbf{b}_0(0)) \\ &+ \sum_{i=1}^k \mathbf{B}_{t_1\tau_i} \left[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i)) \right] \delta t_i \in \mathbf{A}_{\mathbf{x}(t_1)}. \end{aligned} \quad (7.3)$$

Proof. The proof is technically complicated, yet conceptually not very difficult. The idea is to decompose $\mathbf{a}(t, s)$ into several parts, which, after a suitable reparametrisation, form an \mathbf{A} -homotopy. As one may have expected, these parts correspond to "switches" in the needle variation associated with the symbol \mathfrak{w} . Our argument will be therefore inductive w.r.t. k —the number of "switches" in \mathfrak{w} . Formula (7.2) will be obtained from the repetitive usage of Lemma 3.15 for the partial homotopies, and (7.3) will follow from the concrete form of these homotopies. The preservation of the uniform regularity will be obtained using the technical results introduced in Appendix B.1.

Finally, let us explain the role of the number θ . We know from Theorem B.12 that if a solution of the ODE for a fixed initial condition \mathbf{x}_0 is defined on an interval $[t_0, t_1]$, then so are the solutions for initial conditions close enough to \mathbf{x}_0 . Since the base variation $\mathbf{x}(t, s)$ associated with $u_s^{\mathfrak{w}}(t)$ is obtained as a composition of the solutions of (4.2) with perturbations on intervals of length $s\delta t_i$ and $s\delta t$, it is clear that, if numbers δt_i and δt are bounded and τ_i, τ and v_i fixed, for a given $\mathbf{x}_0(s)$, we can chose $\theta > 0$ such that the trajectory $\mathbf{x}(t, s)$ will stay close enough to $\mathbf{x}(t)$ to be well-defined for all $0 \leq s \leq \theta$ and all $[t_0, t_1]$.

In our inductive reasoning it will be more convenient to assume that all the data depends on an additional parameter $p \in P$ (i.e., we have $\mathbf{x}_0(s, p)$ instead of $\mathbf{x}_0(s)$, $\mathbf{a}(t, s, p)$ instead of $\mathbf{a}(t, s)$, etc.). In the assertion we demand that (7.2) and (7.3) hold for each fixed $p \in P$. Moreover, for fixed (τ_i, v_i, τ) we want $\mathbf{d}^{\mathfrak{w}, p}(s)$ to be uniformly regular w.r.t. p , δt_i , and δt at $s = 0$ if $\mathbf{b}_0(s, p)$ is uniformly regular w.r.t. p at $s = 0$.

In what follows we will need two technical lemmas.

Lemma 7.5. *Let $t \mapsto \mathbf{a}(t, s, \tilde{p}) = \mathbf{f}(\mathbf{x}(t, s, \tilde{p}), v(t))$, with $t \in [\tilde{t}_0, \tilde{t}_1]$, be a family of bounded measurable admissible paths over $\mathbf{x}(t, s, \tilde{p})$ parameterised by $\tilde{p} \in \tilde{P}$. Let $s \mapsto \tilde{\mathbf{b}}_0(s, \tilde{p})$ be a family of bounded measurable \mathbf{A} -paths over $\mathbf{x}(t_0, s, \tilde{p})$. There exists a number $\theta > 0$ and a family of bounded measurable \mathbf{A} -paths $s \mapsto \tilde{\mathbf{d}}_1^{\tilde{p}}(s)$ defined for $0 \leq s \leq \theta$ such that*

$$[\tilde{\mathbf{b}}_0(s, \tilde{p})]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(t, \varepsilon, \tilde{p}), v(t))]_{t \in [\tilde{t}_0, \tilde{t}_1]} = [\mathbf{f}(\mathbf{x}(t, 0, p), v(t))]_{t \in [\tilde{t}_0, \tilde{t}_1]} [\tilde{\mathbf{d}}_1^{\tilde{p}}(s)]_{s \in [0, \varepsilon]}, \quad (7.4)$$

for all $\varepsilon \leq \theta$.

Explicitly, $\tilde{\mathbf{d}}_1^{\tilde{p}}(s) = \mathbf{B}_{t\tilde{t}_0}^v \left[\tilde{\mathbf{b}}_0(s, \tilde{p}) \right]$, where $\mathbf{B}_{t\tilde{t}_0}^v$ is a parallel transport operator associated with the control $v(t)$. Moreover, if $\tilde{\mathbf{b}}_0(s, \tilde{p})$ is uniformly regular w.r.t. \tilde{p} at $s = 0$, then so is $\tilde{\mathbf{d}}_1^{\tilde{p}}(s)$.

Proof of the lemma. Fix $\tilde{p} \in \tilde{P}$ and consider $\mathbf{b}(t, s, \tilde{p}) := \mathbf{B}_{t\tilde{t}_0}^v \left[\tilde{\mathbf{b}}_0(s, \tilde{p}) \right]$. It follows from the definition of the operator of parallel transport $\mathbf{B}_{t\tilde{t}_0}^v$ that the pair $(\mathbf{a}(t, s, \tilde{p}), \mathbf{b}(t, s, \tilde{p}))$ is

and \mathbf{A} -homotopy over $\mathbf{x}(t, s, \tilde{p})$ (see Remark 4.5). Now (7.4) follows directly from Lemma 3.15, since $\mathbf{b}(\tilde{t}_1, s, p) = \mathbf{B}_{\tilde{t}_1 \tilde{t}_0}^v [\tilde{\mathbf{b}}_0(s, \tilde{p})] = \mathbf{d}_1^{\tilde{p}}(s)$.

Finally, since by Remark 4.5 the map $\mathbf{B}_{\tilde{t}_1 \tilde{t}_0}^v(\cdot)$ is continuous for every fixed \tilde{t}_1 and \tilde{t}_0 , in light of Lemma B.11, it preserves the uniform regularity of $\tilde{\mathbf{b}}_0(s, \tilde{p})$. \square

The second lemma is the following one.

Lemma 7.6. *Let $t \mapsto \mathbf{a}(t, s, \tilde{p}) = \mathbf{f}(\mathbf{x}(t, s, \tilde{p}), v(t))$ be a family of bounded measurable admissible paths over $\mathbf{x}(t, s, \tilde{p})$ parametrised by $\tilde{p} \in P$. Let $s \mapsto \tilde{\mathbf{b}}_0(s, \tilde{p})$ be a family of bounded measurable \mathbf{A} -paths over $\mathbf{x}(\tilde{t}_0 + sc, s, \tilde{p})$. Then there exists a number $\theta > 0$ and a family of bounded measurable \mathbf{A} -paths $s \mapsto \mathbf{d}_2^{\tilde{p}, c, d}(s)$ defined for $0 \leq s \leq \theta$ such that*

$$[\tilde{\mathbf{b}}_0(s, \tilde{p})]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(\tilde{t}_0 + t, \varepsilon, \tilde{p}), v(t))]_{t \in [\varepsilon c, s \varepsilon d]} = [\mathbf{d}_2^{\tilde{p}, c, d}(s)]_{s \in [0, \varepsilon]}, \quad (7.5)$$

for every $\varepsilon \leq \theta$.

Moreover, if $\tilde{\mathbf{b}}_0(s, \tilde{p})$ is uniformly regular w.r.t. \tilde{p} at $s = 0$ and \tilde{t}_0 is a regular point of the control $v(t)$, then $\mathbf{d}_2^{\tilde{p}, c, d}(s)$ is uniformly regular w.r.t. \tilde{p} , c , and d at $s = 0$. Finally,

$$\mathbf{d}_2^{\tilde{p}, c, d}(0) = \tilde{\mathbf{b}}_0(s, \tilde{p}) + (d - c) \mathbf{f}(\mathbf{x}(\tilde{t}_0, 0, \tilde{p}), v(\tilde{t}_0)). \quad (7.6)$$

Proof of the lemma. For notation simplicity let forget about the \tilde{p} -dependence. Define

$$\widehat{\mathbf{b}}_0(s) := \mathbf{B}_{\tilde{t}_0(\tilde{t}_0 + sc)} [\tilde{\mathbf{b}}_0(s) - c \mathbf{f}(\mathbf{x}(\tilde{t}_0 + cs), v(\tilde{t}_0 + cs))].$$

Clearly, $\widehat{\mathbf{b}}_0(s)$ is an admissible paths over $\mathbf{x}(\tilde{t}_0, s)$. Now define a pair of maps

$$\begin{aligned} \mathbf{a}(t, s) &= s \mathbf{f}(\mathbf{x}(\tilde{t}_0 + ts, s), v(\tilde{t}_0 + ts)), \\ \mathbf{b}(t, s) &= \mathbf{B}_{(\tilde{t}_0 + ts)\tilde{t}_0}^v [\tilde{\mathbf{b}}_0(s)] + t \mathbf{f}(\mathbf{x}(\tilde{t}_0 + ts, s), v(\tilde{t}_0 + ts)), \end{aligned}$$

where $t \in [c, d]$ and $s \in [0, \theta]$. We shall prove that this pair is an \mathbf{A} -homotopy.

If this is the case, then clearly (7.5) follows from Lemma 3.15 since the initial-point \mathbf{A} -homotopy is

$$\begin{aligned} \mathbf{b}(c, s) &= \mathbf{B}_{(\tilde{t}_0 + cs)\tilde{t}_0}^v \mathbf{B}_{\tilde{t}_0(\tilde{t}_0 + sc)}^v [\tilde{\mathbf{b}}_0(s) - c \mathbf{f}(\mathbf{x}(\tilde{t}_0 + sc), v(\tilde{t}_0 + sc))] + \\ &\quad + c \mathbf{f}(\mathbf{x}(\tilde{t}_0 + sc), v(\tilde{t}_0 + sc)) = \tilde{\mathbf{b}}_0(s), \end{aligned}$$

the final-point \mathbf{A} -homotopy is

$$\begin{aligned} \mathbf{d}_2^{c, d}(s) := \mathbf{b}(d, s) &= \mathbf{B}_{(\tilde{t}_0 + ds)\tilde{t}_0}^v [\tilde{\mathbf{b}}_0(s) - c \mathbf{f}(\mathbf{x}(\tilde{t}_0 + sc), v(\tilde{t}_0 + sc))] + \\ &\quad + d \mathbf{f}(\mathbf{x}(\tilde{t}_0 + sd), v(\tilde{t}_0 + sd)), \end{aligned}$$

and, by Lemma 3.17, $[\mathbf{a}(t, s)]_{t \in [c, d]} = [\mathbf{f}(\tilde{t}_0 + t, s), v(\tilde{t}_0 + t)]_{t \in [cs, ds]}$. Evaluating the formula for $\mathbf{d}_2^{c, d}(s)$ at $s = 0$ we get (7.6).

Finally, $\mathbf{d}_2^{c,d}(s)$ is uniformly regular if $\tilde{\mathbf{b}}_0(s)$ is and \mathbf{t}_0 is a regular point of $v(t)$. Indeed, we can use the results from Appendix B.1. The point is to observe that $\mathbf{d}_2^{c,d}(s)$ is obtained from measurable maps $v(\tilde{t}_0 + s)$ and $\mathbf{b}_0(s)$ regular at $s = 0$ by operations described in Propositions B.6–B.9 and Lemmas B.10–B.11 which preserve the uniform regularity. One has also to use the fact that $\mathbf{B}_{t_1 t_0}^v(\mathbf{b}_0)$, $f(x, u)$, and $x(t, s)$ are continuous maps (cf. Remark 4.5).

Now it remains to check that $\mathbf{a}(t, s)$ and $\mathbf{b}(t, s)$ are indeed an \mathbf{A} -homotopy. Let us calculate the WT-derivatives:

$$\begin{aligned} \partial_s \mathbf{a}^i(t, s) &\stackrel{\text{WT}}{=} \mathbf{f}^i(\mathbf{x}(\tilde{t}_0 + ts, s), v(\tilde{t}_0 + ts)) + ts \partial_{\tilde{t}} \mathbf{f}^i(\mathbf{x}(\tilde{t}, s), v(\tilde{t})) \Big|_{\tilde{t}=\tilde{t}_0+ts} + \\ &\quad + s \partial_s \mathbf{f}^i(\mathbf{x}(\tilde{t}, s), v(\tilde{t})) \Big|_{\tilde{t}=\tilde{t}_0+ts}, \end{aligned}$$

and

$$\begin{aligned} \partial_t \mathbf{b}^i(t, s) &\stackrel{\text{WT}}{=} \partial_{\tilde{t}} \mathbf{B}_{\tilde{t} t_0}^v \left[\widehat{\mathbf{b}}_0(s) \right]^i \Big|_{\tilde{t}=\tilde{t}_0+ts} + \mathbf{f}^i(\mathbf{x}(\tilde{t}_0 + ts, s), v(\tilde{t}_0 + ts)) + \\ &\quad + ts \partial_{\tilde{t}} \mathbf{f}^i(\mathbf{x}(\tilde{t}, s), v(\tilde{t})) \Big|_{\tilde{t}=\tilde{t}_0+ts}. \end{aligned}$$

Now, since $\mathbf{f}(\mathbf{x}(\tilde{t}, s), v(\tilde{t}))$ and $\mathbf{B}_{\tilde{t} t_0}^v \left[\widehat{\mathbf{b}}_0(s) \right]$ is an \mathbf{A} -homotopy (cf. Lemma 7.5), we have

$$\partial_{\tilde{t}} \mathbf{B}_{\tilde{t} t_0}^v \left[\widehat{\mathbf{b}}_0(s) \right]^i - \partial_s \mathbf{f}^i(\mathbf{x}(\tilde{t}, s), v(\tilde{t})) \stackrel{\text{WT}}{=} c_{jk}^i(\mathbf{x}(\tilde{t}, s)) \mathbf{B}_{\tilde{t} t_0}^v \left[\widehat{\mathbf{b}}_0(s) \right]^j \mathbf{f}^k(\mathbf{x}(\tilde{t}, s), v(\tilde{t})).$$

Consequently,

$$\begin{aligned} \partial_t \mathbf{b}^i(t, s) - \partial_s \mathbf{a}^i(t, s) &\stackrel{\text{WT}}{=} s \left[\partial_{\tilde{t}} \mathbf{B}_{\tilde{t} t_0}^v \left[\widehat{\mathbf{b}}_0(s) \right]^i - \partial_s \mathbf{f}^i(\mathbf{x}(\tilde{t}, s), v(\tilde{t})) \right] \Big|_{\tilde{t}=\tilde{t}_0+ts} \\ &\stackrel{\text{WT}}{=} s c_{jk}^i(\mathbf{x}(\tilde{t}, s)) \mathbf{B}_{\tilde{t} t_0}^v \left[\widehat{\mathbf{b}}_0(s) \right]^j \mathbf{f}^k(\mathbf{x}(\tilde{t}, s), v(\tilde{t})) \Big|_{\tilde{t}=\tilde{t}_0+ts} \\ &= c_{jk}^i(\mathbf{x}(\tilde{t}_0 + ts, s)) \mathbf{b}^j(t, s) \mathbf{a}^k(t, s). \end{aligned}$$

□

Now we return to the inductive proof of Theorem 7.4. We will prove first that the assertion is true for $t_1 = \tau$. Our argument will be inductive w.r.t. k —the number of switches in the symbol $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$.

Step 1, $k = 0$. We start with $k = 0$. This means that $u_s^{\mathbf{w}}(t) = u(t)$. For $\delta t = 0$ we simply have $\mathbf{a}(t, s) = \mathbf{f}(\mathbf{x}(t, s, p), u(t))$ with $t \in [t_0, \tau]$, where $\mathbf{x}(t_0, s, p) = \mathbf{x}_0(s, p)$. Now we can use the Lemma 7.5 taking $\tilde{t}_0 = t_0$, $\tilde{t}_1 = \tau$, $\tilde{p} = p$, $\tilde{\mathbf{b}}_0(s, \tilde{p}) = \mathbf{b}_0(s, p)$, and $v(t) = u(t)$ to get the assertion.

If $\delta t \neq 0$, things are a little more complicated. We have $\mathbf{a}(t, s, p) = \mathbf{f}(\mathbf{x}(t, s, p), u(t))$ where $t \in [t_0, \tau + s\delta t]$ and $\mathbf{x}(t_0, s, p) = \mathbf{x}_0(s, p)$. We can decompose

$$\begin{aligned} &[\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u(t))]_{t \in [t_0, \tau + \varepsilon \delta t]} \\ &= [\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u(t))]_{t \in [t_0, \tau]} \cdot [\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u(t))]_{t \in [\tau, \tau + \varepsilon \delta t]}. \end{aligned} \tag{7.7}$$

Now using the assertion for $\delta t = 0$ we get

$$[\mathbf{b}_0(s, p)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u(t))]_{t \in [t_0, \tau]} = [\mathbf{f}(\mathbf{x}(t, 0, p), u(t))]_{t \in [t_0, \tau]} [\mathbf{d}_1^p(s)]_{s \in [0, \varepsilon]}, \quad (7.8)$$

where $\mathbf{d}_1^p(s)$ is uniformly regular w.r.t. p at $s = 0$, and $\mathbf{d}_1^p(0) = \mathbf{B}_{\tau t_0}[\mathbf{b}_0(0, p)]$. Next, using Lemma 7.6 for $\tilde{t}_0 = \tau$, $c = 0$, $d = \delta t$, $\tilde{p} = p$, $\tilde{\mathbf{b}}_0(s, \tilde{p}) = \mathbf{d}_1^p(s)$, and $v(t) = u(t)$, we get

$$[\mathbf{d}_1^p(s)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(\tau + t, \varepsilon, \tilde{p}), v(\tau + t))]_{t \in [0, \varepsilon \delta t]} = [\mathbf{d}_2^{p, \delta t}(s)]_{s \in [0, \varepsilon]}, \quad (7.9)$$

where $\mathbf{d}_2^{p, \delta t}(s)$ is uniformly regular w.r.t. p , and δt at $s = 0$ and

$$\mathbf{d}_2^{p, \delta t}(0) = \mathbf{d}_1^p(0) + \delta t \mathbf{f}(\mathbf{x}(\tau, 0, p), u(\tau)).$$

Multiplying (7.8) by $[\mathbf{f}(\mathbf{x}(\tau + t, \varepsilon, \tilde{p}), v(\tau + t))]_{t \in [0, \varepsilon \delta t]}$, using (7.7) and (7.9), and taking $\mathbf{d}^{\mathbf{w}, p}(s) := \mathbf{d}_2^{p, \delta t}(s)$, we get the assertion.

Step 2. Assume that the assertion holds for all $l < k$. Consider a symbol $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$. We will distinguish the following two situations:

Situation 2.A. Not all τ_i are equal. In particular,

$$t_0 < \tau_1 \leq \dots \leq \tau_l < \tau_{l+1} \leq \dots \leq \tau_k \leq \tau$$

for some $l < k$. We can now use the inductive assumption for a symbol $\mathbf{w}_1 = (\tau_i, v_i, \tau = \tau_l, \delta t_i, \delta t = 0)_{i=1, \dots, l}$ to get

$$\begin{aligned} [\mathbf{b}_0(s, p)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u_\varepsilon^{\mathbf{w}_1}(t))]_{t \in [t_0, \tau_l]} \\ = [\mathbf{f}(\mathbf{x}(t, 0, \varepsilon), u(t))]_{t \in [t_0, \tau_l]} [\mathbf{d}^{\mathbf{w}_1, p}(s)]_{s \in [0, \varepsilon]}, \end{aligned} \quad (7.10)$$

where $\mathbf{d}^{\mathbf{w}_1, p}(s)$ is uniformly regular w.r.t. $p, \delta t_1, \dots, \delta t_l$ at $s = 0$, and

$$\mathbf{d}^{\mathbf{w}_1, p}(0) = \mathbf{B}_{\tau t_0}(\mathbf{b}_0(0, p)) + \sum_{i=1}^l \mathbf{B}_{\tau \tau_i} \left[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i)) \right] \delta t_i.$$

Using the inductive assumption for $\mathbf{w}_2 = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=l+1, \dots, k}$ with $t_0 = \tau_l$ and $\mathbf{b}_0(s, p, \delta t_1, \dots, \delta t_l) = \mathbf{d}_1^{\mathbf{w}_1, p}(s)$, we get

$$\begin{aligned} [\mathbf{d}^{\mathbf{w}_1, p}(s)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u_\varepsilon^{\mathbf{w}}(t))]_{t \in [\tau_l, \tau]} \\ = [\mathbf{f}(\mathbf{x}(t, 0, \varepsilon), u(t))]_{t \in [\tau_l, \tau]} [\mathbf{d}^{\mathbf{w}_1, \mathbf{w}_2, p}(s)]_{s \in [0, \varepsilon]}, \end{aligned} \quad (7.11)$$

where $\mathbf{d}^{\mathbf{w}_1, \mathbf{w}_2, p}(s)$ is uniformly regular w.r.t. $p, \delta t_i$ and δt at $s = 0$, and

$$\begin{aligned} \mathbf{d}^{\mathbf{w}_1, \mathbf{w}_2, p}(0) &= \mathbf{B}_{\tau \tau_l}(\mathbf{d}^{\mathbf{w}_1, p}(0)) + \sum_{i=l}^k \mathbf{B}_{\tau \tau_i} \left[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i)) \right] \delta t_i \\ &= \mathbf{B}_{\tau t_0}(\mathbf{b}_0(0, p)) + \sum_{i=1}^k \mathbf{B}_{\tau \tau_i} \left[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i)) \right] \delta t_i. \end{aligned}$$

Multiplying (7.10) by $[\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u_\varepsilon^{\mathbf{w}}(t))]_{t \in [\tau_i, \tau]}$, using (7.11), and taking $\mathbf{d}^{\mathbf{w}, p}(s) := \mathbf{d}^{\mathbf{w}_1, \mathbf{w}_2, p}(s)$, we get the assertion.

Situation 2.B. If all τ_i are equal then either

$$\tau_1 = \dots \tau_k = \tau \quad \text{or} \quad \tau_1 = \dots \tau_k < \tau.$$

2.B.1. In the first case using the assertion for $\mathbf{w}_1 = (\tau_i, v_i, \delta t_i, \tau, \delta t - \delta t_k)_{i=1,2,\dots,k-1}$ we get

$$\begin{aligned} & [\mathbf{b}_0(t, s)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(t, \varepsilon, p), u_\varepsilon^{\mathbf{w}_1}(t))]_{t \in [t_0, \tau + (\delta t - \delta t_k) \varepsilon]} \\ &= [\mathbf{f}(\mathbf{x}(t, 0, p), u(t))]_{t \in [t_0, \tau]} [\mathbf{d}_1^{\mathbf{w}_1, p}(s)]_{s \in [0, \varepsilon]}, \end{aligned} \quad (7.12)$$

where $\mathbf{d}_1^{\mathbf{w}_1, p}(s)$ is uniformly regular w.r.t. $p, \delta t_1, \dots, \delta t_{k-1}, \delta t - \delta t_k$ at $s = 0$ and

$$\mathbf{d}_1^{\mathbf{w}_1, p}(0) = \mathbf{B}_{\tau t_0}(\mathbf{b}_0(0, p)) + (\delta t - \delta t_k) \mathbf{f}(\mathbf{x}(\tau), u(\tau)) + \sum_{i=1}^{k+1} \mathbf{B}_{\tau \tau_i} [\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i))] \delta t_i.$$

Now using Lemma 7.6 for $\tilde{t}_0 = \tau$, $c = \delta t - \delta t_k$, $\tilde{v}(t) = v_k$, $\tilde{p} = (p, \delta t_1, \dots, \delta t_{k-1}, \delta t - \delta t_k)$, and $\tilde{\mathbf{b}}_0(s, \tilde{p}) = \mathbf{d}_1^{\mathbf{w}_1, p}(s)$ we get

$$[\mathbf{d}_1^{\mathbf{w}_1, p}(s)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(\tau + t, \varepsilon, p), v_k)]_{t \in [\varepsilon(\delta t - \delta t_{k-1}), \varepsilon \delta t]} = [\mathbf{d}_2^{\mathbf{w}_1, \delta t, \delta t_k, p}(s)]_{s \in [0, \varepsilon]}, \quad (7.13)$$

where $\mathbf{d}_2^{\mathbf{w}_1, \delta t, \delta t_k, p}(s)$ is uniformly regular w.r.t. $p, \delta t_1, \dots, \delta t_k, \delta t$ at $s = 0$ and

$$\mathbf{d}_2^{\mathbf{w}_1, \delta t, \delta t_k, p}(0) = \mathbf{d}_1^{\mathbf{w}_1, p}(0) - \delta t_k \mathbf{f}(\mathbf{x}(\tau, 0, p), u(\tau)).$$

Again multiplying (7.11) by $[\mathbf{f}(\mathbf{x}(\tau + t, \varepsilon, p), v_k)]_{t \in [\varepsilon(\delta t - \delta t_{k-1}), \varepsilon \delta t]}$, using (7.13) and taking $\mathbf{d}^{\mathbf{w}, p}(s) = \mathbf{d}_2^{\mathbf{w}_1, \delta t, \delta t_k, p}(s)$ we get the assertion.

2.B.1. If $\tau_1 = \dots = \tau_k < \tau$ we can use the result from 2.B.1 for a symbol $\mathbf{w}_1 = (\tau_i, v_i, \delta t_i, \tau = \tau_k, \delta t = 0)_{i=1,\dots,k-1}$ and then use the inductive assumption for $[t_0, \tau] = [\tau_k, \tau]$ and $\mathbf{w}_2 = (\tau, \delta t)$ on $[t_0 = \tau_k, \tau]$ in essentially the same way as in the case A. The inductive argument is now complete.

Finally, to obtain the assertion for t_1 not τ , one has just to proceed as in the step 1 with $\delta t = 0$ and use Lemma 7.5 again, taking $\tilde{t}_0 = \tau$, $\tilde{t}_1 = t_1$, $\tilde{\mathbf{b}}_0(s, \tilde{p})$ to be the final-point \mathbf{A} -homotopy $\mathbf{d}^{\mathbf{w}, p}(s)$ derived for $t_1 = \tau$, and the control $v(t) = u(t) = u_s^{\mathbf{w}}(t + s \delta t)$. \square

The set of infinitesimal variations K_τ^u

Remark 7.7. Observe that choosing $\mathbf{b}_0(s) \equiv \theta_{\mathbf{x}_0}$ in Theorem 7.4 we obtain an admissible path $\mathbf{d}^{\mathbf{w}}(s)$, regular at $s = 0$, defined for $0 \leq s < \theta$, and satisfying

$$[\mathbf{a}(t, \varepsilon)]_{t \in [t_0, t_1 + \varepsilon \delta t]} = [\mathbf{a}(t)]_{t \in [t_0, t_1]} [\mathbf{d}^{\mathbf{w}}(s)]_{s \in [0, \varepsilon]}, \quad (7.14)$$

for $\varepsilon \leq \theta$.

We define the set \mathbf{K}_τ^u consisting of elements of the form $\mathbf{d}^{\mathbf{w}}(0)$, where $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$ are symbols with τ fixed:

$$\mathbf{K}_\tau^u := \left\{ \mathbf{B}_{t_1\tau}[\mathbf{f}(\mathbf{x}(\tau), u(\tau))]\delta t + \sum_{i=1}^k \mathbf{B}_{t_1\tau_i} \left[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i)) \right] \delta t_i : (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k} \text{ is a symbol} \right\} \subset \mathbf{A}_{\mathbf{x}(t_1)}.$$

We will call \mathbf{K}_τ^u the *set of infinitesimal variations of the trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$ associated with the regular $\tau \in (t_0, t_1)$* .

The set \mathbf{K}_τ^u can be interpreted as the set of all generalised directions in $\mathbf{A}_{\mathbf{x}(t_1)}$ in which one can move the final base-point $\mathbf{x}(t_1)$ by performing needle variations, associated with symbols $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$ with fixed τ and trivial initial base-point variations $\mathbf{x}_0(s) \equiv \mathbf{x}_0$.

The geometry of \mathbf{K}_τ^u will be an object of our main interests in Chapter 9. Now let us note the following property

Lemma 7.8. *The set \mathbf{K}_τ^u is a convex cone in $\mathbf{A}_{\mathbf{x}(t_1)}$.*

Proof. Take symbols $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$, $\mathbf{w}' = (\tau'_i, v'_i, \tau, \delta t'_i, \delta t')_{i=1, \dots, k'}$ and numbers $\nu, \nu' \geq 0$. We have to find a symbol \mathbf{v} such that

$$\mathbf{d}^{\mathbf{v}}(0) = \nu \mathbf{d}^{\mathbf{w}}(0) + \nu' \mathbf{d}^{\mathbf{w}'}(0).$$

Due to Remark 7.3, we may change the symbol by adding $(\tau_i, v_i, \delta t_i = 0)$ without changing the variation $u_\varepsilon^{\mathbf{w}}$. As we see from the form of (7.3), such an addition will not change $\mathbf{d}^{\mathbf{w}}(0)$. Consequently, we may assume that $k = k'$, $\tau_i = \tau'_i$, $v_i = v'_i$, and the symbols \mathbf{w} and \mathbf{w}' differ only by δt_i and δt . Now consider the symbol $\mathbf{v} = (\tau_i, v_i, \tau, \nu \delta t_i + \nu' \delta t'_i, \nu \delta t + \nu' \delta t')_{i=1, \dots, k}$. The formula (7.3) (for $\mathbf{b}_0(0, p) = 0$) is linear with respect to δt_i and δt , hence

$$\mathbf{d}^{\mathbf{v}}(0) = \nu \mathbf{d}^{\mathbf{w}}(0) + \nu' \mathbf{d}^{\mathbf{w}'}(0).$$

□

At the end of this chapter we define several geometric objects which will play an important role in Chapter 9.

Consider the real line \mathbb{R} with the canonical coordinate $t \in \mathbb{R}$. The tangent space $\mathbf{T}_{\underline{x}}\mathbb{R}$ is spanned by the canonical vector ∂_t . Denote by $\Lambda_{\underline{x}}$ the ray

$$\Lambda_{\underline{x}} := \mathbb{R}_+ \cdot (-\partial_t) \subset \mathbf{T}_{\underline{x}}\mathbb{R},$$

and by $\Lambda_{\mathbf{x}}$ the ray

$$\Lambda_{\mathbf{x}} := \theta_x \oplus \Lambda_{\underline{x}} \subset E_x \oplus \mathbf{T}_{\underline{x}}\mathbb{R} = \mathbf{A}_{\mathbf{x}},$$

where $\mathbf{x} = (x, \underline{x})$. Finally define

$$\mathcal{K}_\tau^u := \text{conv} \left\{ \mathbf{B}_{t_1 t_0}(\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)}), \mathbf{K}_\tau^u \right\},$$

where $\mathcal{S}_0 = \Phi_0(\mathbf{T}_{z_0} S_0)$ was defined in Chapter 5. By Theorem 7.4, \mathcal{K}_τ^u has an interpretation of the set of all generalised directions in $\mathbf{A}_{\mathbf{x}(t_1)}$ in which one can move the final base-point $\mathbf{x}(t_1)$ by performing needle variations, associated with symbols $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$ with fixed τ and initial base-point variations in the directions of $\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)}$.

Technical lemmas

In this chapter we prove two technical results about E -homotopies — Lemmas 8.1 and 8.2, which will be crucial in the proof of Theorems 5.1 and 5.4. To discuss briefly the results, given a family of smooth curves $x_{\vec{r}} : I \rightarrow \mathbb{R}^m$, parameterized by $\vec{r} \in B^m(0, 1) \subset \mathbb{R}^m$, which emerges from a single point $x_{\vec{r}}(0) = 0$ and points into every direction $\dot{x}_{\vec{r}}(0) = \vec{r}$, it is quite obvious that, for every $t > 0$ small enough, there exists a curve $x_{\vec{r}_0}$ from this family which reaches 0 at time t . A similar result holds for families of admissible curves on a skew-algebroid E . Any such family which is sufficiently regular and emerges from a single point into every direction in E will realise a zero homotopy class. This is Lemma 8.1. In Lemma 8.2 we prove that two sufficiently regular families of admissible paths in E emerging from a single point must have a nonempty intersection of homotopy classes, provided that the set of their initial (generalized) velocities is rich enough.

These results seem to be quite natural and they are indeed, if such an algebroid is (locally) integrable. In this case E -homotopy classes can be represented by points on a finite-dimensional manifold. However, if E is not integrable, E -homotopy is just a relation in the space of bounded measurable curves. Therefore to prove the results we have to pass through the Banach space setting. The main idea in the proof is to semi-parametrise the E -homotopy classes by a finite dimensional-space and reduce the problem to a finite-dimensional topological problem. By a semi-parametrisation we mean an epimorphism from a finite-dimensional space to the space of E -homotopy classes.

Local coordinates

Since we are going to work in a Banach space setting it is convenient to introduce local coordinates on an algebroid E . Consider coordinates $(x^a, y^i) \in U \times \mathbb{R}^m \subset \mathbb{R}^n \times \mathbb{R}^m$ trivialising the bundle $\tau : E \rightarrow M$ around a point $p \in M$. We may assume that p corresponds to $0 \in U$. As usual, we will denote the structural functions of E in these coordinates by $\rho_i^a(x)$ and $c_{jk}^i(x)$. Since these functions are smooth, we can assume (after restricting ourselves to a compact neighborhood $\bar{V} \ni 0$ in \mathbb{R}^n) that they are bounded (by numbers C_ρ and C_c , respectively) and Lipschitz w.r.t. x^a (with constants L_ρ and L_c , respectively). It will be convenient to think of $V \times \mathbb{R}^m$ with those functions as of a (local) AL algebroid. Observe that every bounded measurable E -path with the base initial-point p is represented by a pair of paths $(x(t), a(t)) \in \mathbb{R}^n \times \mathbb{R}^m$, where $a(t)$ is bounded measurable

and $x(t)$ is an AC-solution of the ODE

$$\begin{cases} \dot{x}^b(t) = \rho_i^b(x(t))a^i(t), \\ x^b(0) = 0. \end{cases}$$

As we see, $x(t)$ is determined entirely by $a(t)$. We can thus identify the space $\mathcal{ADM}_p(I, E)$ of bounded measurable admissible paths originated at p with the space $\mathcal{BM}(I, \mathbb{R}^m)$ of bounded measurable maps $a : I \rightarrow \mathbb{R}^m$. We will consequently speak of algebroid homotopy classes in $\mathcal{BM}(I, \mathbb{R}^m)$. Note that $\mathcal{BM}(I, \mathbb{R}^m)$, equipped with the L^1 -norm, is a Banach space. We will denote this norm simply by $\|\cdot\|$. The same symbol will be also used for L^1 -norm in \mathbb{R}^m . In our considerations we will understand a product of Banach spaces $(\mathcal{B}_1, \|\cdot\|_1)$ and $(\mathcal{B}_2, \|\cdot\|_2)$ as a space $\mathcal{B}_1 \times \mathcal{B}_2$ equipped with the norm $\|\cdot\| = \|\cdot\|_1 + \|\cdot\|_2$.

First lemma

Lemma 8.1. *Let $a_{\vec{r}}(\cdot) \in \mathcal{BM}(I, \mathbb{R}^m)$, where $\vec{r} \in B^m(0, 1)$, be a family of E -paths uniformly regular at $t = 0$ w.r.t. \vec{r} and such that $a_{\vec{r}}(0) = \vec{r}$. Then there exists a number $\eta > 0$ with the following property. For every $0 < \varepsilon < \eta$ there exists a vector \vec{r}_0 such that the curve $a_{\vec{r}_0}(t)$, after restricting to the interval $[0, \varepsilon]$, is null- E -homotopic:*

$$[a_{\vec{r}_0}(t)]_{t \in [0, \varepsilon]} = [0].$$

Let us briefly sketch the strategy of the proof. Denote by $c_{\vec{r}}$ a constant path $c_{\vec{r}}(s) = \vec{r}$ in \mathbb{R}^m . We will construct a continuous and invertible (local) map of Banach spaces $\Phi : \mathbb{R}^m \times \mathcal{B} \rightarrow \mathcal{BM}(I, \mathbb{R}^m)$ (the space \mathcal{B} will be specified later) which will have an additional property that the homotopy class of the image is determined by the first factor only

$$[\Phi(\vec{r}, d)] = [c_{\vec{r}}].$$

In such a way we will realise our idea from the introduction to this chapter — \mathbb{R}^m will semi-parametrise all local E -homotopy classes of $\mathcal{ADM}_p(I, E)$. Next, using the map $a : \vec{r} \mapsto a_{\vec{r}}$, we will construct a continuous map of finite-dimensional spaces

$$\mathbb{R}^m \supset B^m(0, 1) \xrightarrow{a} \mathcal{BM}([0, 1], \mathbb{R}^m) \xrightarrow{\Phi^{-1}} \mathbb{R}^m \times \mathcal{B} \xrightarrow{\text{pr}_1} \mathbb{R}^m.$$

A topological argument will prove that 0 lies in the image of this map, hence

$$[a_{\vec{r}_0}] = [c_0] = [0] \quad \text{for some } \vec{r}_0.$$

Proof. Consider an E -path with a constant \mathbb{R}^m -part $a(t) = \vec{r}$, where \vec{r} is a fixed element in \mathbb{R}^m , $t \in [0, 1]$. The associated base path $x(t) \in \mathbb{R}^m$ is the solution of

$$\begin{cases} \partial_t x^b(t) = \rho_i^b(x(t))a^i(t) = \rho_i^b(x(t))r^i, \\ x^b(0) = 0. \end{cases} \quad (8.1)$$

Clearly, if $\|\vec{r}\|$ is small enough, the solution of this equation exists for $t \in [0, 1]$ and is contained entirely in $\overline{V} \subset \mathbb{R}^n$. Now for $a(t)$ and $x(t)$ as above and fixed paths $d \in \mathcal{BM}(I, \mathbb{R}^m)$, $b \in \mathcal{AC}(I, \mathbb{R}^m)$, consider the following system of differential equations:

$$\begin{cases} \partial_s a^i(t, s) = d^i(t) + c_{jk}^i(x(t, s))a^j(t, s)b^k(t), \\ a^i(t, 0) = a^i(t) = r^i, \\ \partial_s x^b(t, s) = \rho_i^b(x(t, s))b^i(t), \\ x^b(t, 0) = x^b(t). \end{cases} \quad (8.2)$$

The existence and regularity of the solutions of (8.2) can be discussed using the theory developed in Appendix B.2. Let us concentrate first on the equation for $x(t, s)$. The right hand-side of this equation is AC in t and Lipschitz in x , the initial value depends AC on a parameter t , and hence, by the standard theory of ODEs, the solution $x(t, s)$ is defined locally and is AC w.r.t. both variables. By shrinking the norm $\|b\|_{\text{sup}}$ we may change the Lipschitz constant in the defining equation. Consequently, for $\|b\|_{\text{sup}}$ (and $\|\vec{r}\|$) small enough, the solution $x(t, s)$ is defined for all $t, s \in [0, 1]$ and entirely contained in \overline{V} .

Now the right hand-side of the first equation in (8.2) is locally Lipschitz w.r.t. a and bounded measurable w.r.t. the parameter t . The initial value $a^i(t, 0)$ depends continuously on t , hence, by Theorem B.13, the solution $a(t, s)$ locally exists, is AC w.r.t. s , and is bounded measurable w.r.t. t . Again, shrinking $\|b\|_{\text{sup}}$ makes the Lipschitz constant smaller, hence for $\|b\|_{\text{sup}}$ small enough $a(t, s)$ is defined for all $t, s \in [0, 1]$.

Now let us consider (8.2) with $b(t) = \int_0^t d(s)ds$, where d is chosen in such a way, that $b(0) = b(1) = 0$. We have

$$\begin{cases} \partial_s a^i(t, s) = \partial_t b^i(t) + c_{jk}^i(x(t, s))a^j(t, s)b^k(t), \\ \partial_s x^b(t, s) = \rho_i^b(x(t, s))b^i(t). \end{cases} \quad (8.3)$$

We recognise equations (3.2) for E -homotopy. Indeed, in such a situation $a(t, s)$ and $b(t, s) = b(t)$ form an E -homotopy with fixed end-points (since $b(0) = b(1) = 0$). Consequently, the homotopy classes of $a(t, 0) = c_{\vec{r}}(t)$ and $a(t, 1)$ are equal. Since $\|b\|_{\text{sup}} \leq \|d\|$, for $\|d\|$ and $\|\vec{r}\|$ small enough, this homotopy is defined for all $t, s \in [0, 1]$. For $a(t)$ and $d(t)$ as above we define

$$\Phi(\vec{r}, d)(t) := a(t, 1).$$

This is a (local) map of Banach spaces

$$\Phi : \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m) \supset W_0 \longrightarrow \mathcal{BM}(I, \mathbb{R}^m),$$

where $\mathcal{BM}_0(I, \mathbb{R}^m) = \{d \in \mathcal{BM}(I, \mathbb{R}^m) : \int_0^1 d(s)ds = 0\}$ is a Banach subspace of $\mathcal{BM}(I, \mathbb{R}^m)$ and W_0 is some open neighbourhood of the point $(0, 0)$. We shall now prove the following:

- (A) Φ maps $(\vec{r}, 0)$ into a constant path $c_{\vec{r}} \in \mathcal{BM}(I, \mathbb{R}^m)$.
- (B) Φ is a continuous map of Banach spaces.

(C) The E -homotopy class of the curve $\Phi(\vec{r}, d)$ is determined by \vec{r} ; that is,

$$[\Phi(\vec{r}, d)] = [\Phi(\vec{r}, 0)] \stackrel{(A)}{=} [c_{\vec{r}}].$$

(D) The map $\tilde{\Phi}(\vec{r}, d) := \Phi(\vec{r}, d) - (c_{\vec{r}} + d)$ is Lipschitz with constant $\frac{1}{6}$.

(E) The map Φ posses a continuous inverse Φ^{-1} defined on some open neighbourhood $V_0 \ni 0$ in $\mathcal{BM}(I, \mathbb{R}^m)$. Moreover, Φ^{-1} is Lipschitz with constant 6.

Property (C) is clear from the construction of Φ , as $a(t, 1) = \Phi(\vec{r}, d)(t)$ and $a(t, 0) = c_{\vec{r}}(t)$ are E -homotopic.

Property (A) is obvious, since $\Phi(\vec{r}, 0)$ is the solution (taken at $s = 1$) of the differential equation $\partial_s a(t, s) = 0$ with the initial condition $a(t, 0) = \vec{r}$.

Property (B) will follow from (D). Indeed, if $\tilde{\Phi}$ is Lipschitz, then $\Phi(\vec{r}, d) = \tilde{\Phi}(\vec{r}, d) + c_{\vec{r}} + d$ is continuous as a sum of continuous maps.

Assuming (D) again, we will be able to prove (E). As one might have expected, the existence and the Lipschitz condition for Φ^{-1} will be proven essentially in the same way as in the standard proof of the inverse function theorem (cf. [Lang, 1985]). First, we will establish a pair of linear isomorphism between Banach spaces

$$\begin{aligned} \mathcal{BM}(I, \mathbb{R}^m) &\xrightarrow{\alpha} \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m), \\ a(t) &\longmapsto \left(\int_0^1 a(s) ds, a(t) - \int_0^1 a(s) ds \right), \end{aligned}$$

and

$$\begin{aligned} \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m) &\xrightarrow{\beta} \mathcal{BM}(I, \mathbb{R}^m), \\ (\vec{r}, d) &\longmapsto c_{\vec{r}} + d. \end{aligned}$$

It is straightforward to verify that α and β are continuous inverses of each other and that α is Lipschitz with constant 3. The map Φ is defined on some open neighbourhood $W_0 \ni (0, 0)$. Take R such that $B(0, 2R) \subset W_0$. The map $\tilde{\Phi}$ is Lipschitz with constant $\frac{1}{6}$ and α is Lipschitz with constant 3; hence $\alpha \circ \tilde{\Phi}$ is Lipschitz with constant $\frac{1}{2}$ and, since it preserves the origin, it maps the ball $B(0, 2R)$ into the ball $B(0, R)$.

Fix now any $a \in \mathcal{BM}(I, \mathbb{R}^m)$ such that $\|a\| < \frac{R}{3}$. We shall construct a unique element $(\vec{r}, d) \in B(0, 2R) \subset W_0$ satisfying $\Phi(\vec{r}, d) = a$. Consider a map $\Phi_a(\vec{r}, d) := \alpha(a - \tilde{\Phi}(\vec{r}, d))$. From the Lipschitzity of α and $\tilde{\Phi}$ we deduce that

$$\|\Phi_a(\vec{r}, d)\| \leq 3\|a - \tilde{\Phi}(\vec{r}, d)\| \leq 3\|a\| + 3\|\tilde{\Phi}(\vec{r}, d)\| \leq 3 \cdot \frac{R}{3} + \frac{1}{2}\|(\vec{r}, d)\|.$$

Consequently, Φ_a maps the ball $B(0, 2R)$ into $B(0, 2R)$. Moreover,

$$\begin{aligned} \|\Phi_a(\vec{r}, d) - \Phi_a(\vec{r}', d')\| &= \|\alpha(a - \tilde{\Phi}(\vec{r}, d)) - \alpha(a - \tilde{\Phi}(\vec{r}', d'))\| \leq \\ &\leq 3\|\tilde{\Phi}(\vec{r}, d) - \tilde{\Phi}(\vec{r}', d')\| \leq 3 \cdot \frac{1}{6}\|(\vec{r}, d) - (\vec{r}', d')\|, \end{aligned}$$

hence Φ_a is a contraction. Now, using the Banach fixed point theorem, we deduce that Φ_a has a unique fixed point $(\vec{r}, d) \in B(0, 2R)$. Consequently,

$$a - \tilde{\Phi}(\vec{r}, d) = \beta \circ \alpha \left(a - \tilde{\Phi}(\vec{r}, d) \right) = \beta \circ \Phi_a(\vec{r}, d) = \beta(\vec{r}, d) = c_{\vec{r}} + d,$$

and hence $a = \tilde{\Phi}(\vec{r}, d) + c_{\vec{r}} + d = \Phi(\vec{r}, d)$. We have proven the existence of Φ^{-1} .

Take now $a, a' \in \mathcal{BM}(I, \mathbb{R}^m)$, and let $\Phi^{-1}(a) = (\vec{r}, d)$, $\Phi^{-1}(a') = (\vec{r}', d')$. Using the Lipschitz condition for α and $\tilde{\Phi}$ once more, we get

$$\begin{aligned} \|(\vec{r}, d) - (\vec{r}', d')\| &= \|\Phi_a(\vec{r}, d) - \Phi_{a'}(\vec{r}', d')\| = \|\alpha\left(a - \tilde{\Phi}(\vec{r}, d)\right) - \alpha\left(a' - \tilde{\Phi}(\vec{r}', d')\right)\| \leq \\ &\leq 3\|a - a'\| + 3\|\tilde{\Phi}(\vec{r}, d) - \tilde{\Phi}(\vec{r}', d')\| \leq 3\|a - a'\| + 3 \cdot \frac{1}{6}\|(\vec{r}, d) - (\vec{r}', d')\|. \end{aligned}$$

We finish the proof of property **(E)** concluding that

$$\|\Phi^{-1}(a) - \Phi^{-1}(a')\| = \|(\vec{r}, d) - (\vec{r}', d')\| \leq 6\|a - a'\|.$$

We are now left with the proof of **(D)**. This will be done by introducing several integral estimations. In our calculations we will, for simplicity, omit the indices (hence c will stand for c_{jk}^i , a for a^i , etc.). Take pairs (r, d) and (r', d') from $\mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m)$. Denote by $x(t, s)$, $a(t, s)$, $a(t)$, $b(t)$ and $x'(t, s)$, $a'(t, s)$, $a'(t)$, $b'(t)$, respectively, the objects defined as in the construction of Φ for pairs (\vec{r}, d) and (\vec{r}', d') . To begin with, observe that, since $b(t) = \int_0^t d(s)ds$, we have $|b(t)| \leq \int_0^1 |d(s)|ds$; hence

$$\|b\|_{\sup} \leq \|d\|.$$

Similarly, $\|b'\|_{\sup} \leq \|d'\|$ and $\|b - b'\|_{\sup} \leq \|d - d'\|$.

Let us now estimate the difference $|x(t, s) - x'(t, s)|$. Since, by (8.3), $x(t, s) = x(t) + \int_0^s \rho(x(t, \sigma))b(t)d\sigma$, we have

$$\begin{aligned} |x(t, s) - x'(t, s)| &= \left| \int_0^s \left(\rho(x(t, \sigma))b(t) - \rho(x'(t, \sigma))b'(t) \right) d\sigma \right| \leq \\ &\leq \int_0^1 \left| \rho(x(t, s))b(t) - \rho(x'(t, s))b'(t) \right| ds \leq \\ &\leq \int_0^1 \left| \left(\rho(x(t, s)) - \rho(x'(t, s)) \right) b(t) \right| ds + \\ &\quad + \int_0^1 \left| \rho(x'(t, s)) \left(b(t) - b'(t) \right) \right| ds \leq \\ &\leq L_\rho \sup_{(t, s)} |x(t, s) - x'(t, s)| \|d\| + C_\rho \|d - d'\|. \end{aligned}$$

It follows that

$$\sup_{(t, s)} |x(t, s) - x'(t, s)| \left(1 - L_\rho \|d\| \right) \leq C_\rho \|d - d'\|.$$

For $\|d\|$ sufficiently small (in other words, after a possible shrinking of W_0) we will obtain

$$\sup_{(t,s)} |x(t,s) - x'(t,s)| \leq 2C_\rho \|d - d'\|. \quad (8.4)$$

Now introduce

$$\begin{aligned} \delta a(t,s) &:= a(t,s) - (a(t) + sd(t)) \quad \text{and} \\ \delta a'(t,s) &:= a'(t,s) - (a'(t) + sd'(t)). \end{aligned}$$

Note that $\delta a(t,1) = \tilde{\Phi}(\vec{r},d)$ and $\delta a'(t,1) = \tilde{\Phi}(\vec{r}',d')$. From (8.2) we deduce that

$$\delta a(t,s) = \int_0^s c(x(t,\sigma))a(t,\sigma)b(t)d\sigma = \int_0^s c(x(t,\sigma))(\delta a(t,\sigma) + a(t) + \sigma d(t))b(t)d\sigma.$$

Further, for a fixed $s \in [0,1]$,

$$\begin{aligned} \|\delta a(\cdot, s)\| &= \int_0^1 |\delta a(t,s)|dt \leq \int_0^1 \int_0^s |c(x(t,\sigma))(\delta a(t,\sigma) + a(t) + \sigma d(t))b(t)|d\sigma dt \leq \\ &\leq \int_0^1 \int_0^1 C_c(|\delta a(t,s)| + \|\vec{r}\| + |d(t)|)\|b\|_{\sup}dsdt \leq \\ &\leq C_c\|b\|_{\sup} \left(\sup_s \|\delta a(\cdot, s)\| + \|\vec{r}\| + \|d\| \right) \leq \\ &\leq C_c\|d\| \left(\sup_s \|\delta a(\cdot, s)\| + \|\vec{r}\| + \|d\| \right) \end{aligned}$$

and we conclude that

$$\sup_s \|\delta a(\cdot, s)\|(1 - C_c\|d\|) \leq C_c(\|\vec{r}\| + \|d\|).$$

Hence, for $\|d\|$ small enough (after possible shrinking of W_0), we get

$$\sup_s \|\delta a(\cdot, s)\| \leq 2C_c(\|\vec{r}\| + \|d\|). \quad (8.5)$$

Finally,

$$\begin{aligned} \|\delta a(\cdot, s) - \delta a'(\cdot, s)\| &= \int_0^1 |\delta a(t,s) - \delta a'(t,s)|dt \leq \\ &\leq \int_0^1 \int_0^s |c(x(t,\sigma))(\delta a(t,\sigma) + a(t) + \sigma d(t))b(t) + \\ &\quad - c(x'(t,\sigma))(\delta a'(t,\sigma) + a'(t) + \sigma d'(t))b'(t)|d\sigma dt \leq \\ &\leq \int_0^1 \int_0^1 |c(x(t,\sigma))\delta a(t,s)b(t) - c(x'(t,s))\delta a'(t,s)b'(t)|dsdt + \\ &\quad + \int_0^1 \int_0^1 |c(x(t,s))a(t)b(t) - c(x'(t,s))a'(t)b'(t)|dsdt + \\ &\quad + \int_0^1 \int_0^1 |c(x(t,s))d(t)b(t) - c(x'(t,s))d'(t)b'(t)|dsdt =: I_1 + I_2 + I_3 \end{aligned}$$

Now we estimate

$$\begin{aligned}
I_1 &\leq \int_0^1 \int_0^1 \left| c(x(t, s)) - c(x'(t, s)) \right| |\delta a(t, s)| |b(t)| dt ds + \\
&\quad + \int_0^1 \int_0^1 \left| c(x'(t, s)) \right| \left| \delta a(t, s) - \delta a'(t, s) \right| |b(t)| dt ds + \\
&\quad + \int_0^1 \int_0^1 \left| c(x'(t, s)) \right| \left| \delta a'(t, s) \right| |b(t) - b'(t)| dt ds \leq \\
&\leq L_c \sup_{(t,s)} |x(t, s) - x'(t, s)| \sup_s \|\delta a(\cdot, s)\| \cdot \|b\|_{\sup} + C_c \sup_s \|\delta a(\cdot, s) - \delta a'(\cdot, s)\| \cdot \|b\|_{\sup} + \\
&\quad + C_c \sup_s \|\delta a'(\cdot, s)\| \cdot \|b - b'\|_{\sup}.
\end{aligned}$$

Finally, using (8.4), (8.5) and $\|b\|_{\sup} \leq \|d\|$, we get

$$\begin{aligned}
I_1 &\leq L_c 2C_\rho \|d - d'\| 2C_c (\|\vec{r}\| + \|d\|) \|d\| + C_c \sup_s \|\delta a(\cdot, s) - \delta a'(\cdot, s)\| \cdot \|d\| + C_c 2C_c (\|\vec{r}'\| \\
&\quad + \|d'\|) \|d - d'\| = C_c \|d\| \sup_s \|\delta a(\cdot, s) - \delta a'(\cdot, s)\| + \|d - d'\| \cdot F_1(\|r\|, \|d\|, \|r'\|, \|d'\|),
\end{aligned}$$

where F_1 converges to 0 when its arguments do. Similar estimations for I_2 and I_3 will give

$$\begin{aligned}
I_2 &\leq (\|r - r'\| + \|d - d'\|) \cdot F_2(\|r\|, \|d\|, \|r'\|, \|d'\|), \\
I_3 &\leq (\|r - r'\| + \|d - d'\|) \cdot F_3(\|r\|, \|d\|, \|r'\|, \|d'\|),
\end{aligned}$$

where F_2 and F_3 behave as F_1 . Putting together the partial results, we would get

$$\sup_s \|\delta a(\cdot, s) - \delta a'(\cdot, s)\| (1 - 2C_c \|d\|) \leq (\|r - r'\| + \|d - d'\|) \cdot F(\|r\|, \|d\|, \|r'\|, \|d'\|),$$

where F converges to 0 when its arguments do. As $\tilde{\Phi}(r, d) = \delta a(t, 1)$ and $\tilde{\Phi}(r', d') = \delta a'(t, 1)$, for W_0 small enough, $\tilde{\Phi}$ is Lipschitz with constant $\frac{1}{6}$. That proves property **(D)**.

Now using properties **(A)**–**(E)** of Φ we will make the final step of the proof of Lemma 8.1. The family $a_{\vec{r}}(t)$ is uniformly regular w.r.t. $\vec{r} \in \overline{B}^m(0, 1)$ at $t = 0$. The family $c_{\vec{r}}(t)$ has the same properties, so $a_{\vec{r}}(t) - c_{\vec{r}}(t)$ is also uniformly regular (cf. Proposition B.8) and

$$\int_0^t |a_{\vec{r}}(s) - c_{\vec{r}}(s)| ds = t |a_{\vec{r}}(0) - c_{\vec{r}}(0)| + f(t, \vec{r}) = t \cdot 0 + f(t, \vec{r}),$$

where $\frac{1}{t} f(t, \vec{r})$ converges uniformly to 0 as $t \rightarrow 0$. Hence, there exists a number $\eta > 0$ such that

$$\int_0^\varepsilon |a_{\vec{r}}(s) - c_{\vec{r}}(s)| ds < \frac{1}{12} \varepsilon,$$

for every $0 \leq \varepsilon \leq \eta$ and $\vec{r} \in \overline{B}^m(0, 1)$. Reparametrising the paths by the rule

$$\tilde{a}_{\vec{r}}(t) := \varepsilon a_{\vec{r}}(\varepsilon t),$$

we will obtain another uniformly regular family of paths satisfying $\tilde{a}_{\vec{r}}(0) = \varepsilon \vec{r}$.

By the uniform regularity of $\tilde{a}_{\vec{r}}$, the map $\tilde{a} : \vec{r} \mapsto \tilde{a}_{\vec{r}}$ is a continuous map from $B^m(0, 1)$ to $\mathcal{BM}(I, \mathbb{R}^m)$ with L_1 -topology. For η small enough \tilde{a} takes values in $\Phi(W_0)$. Composing \tilde{a} with Φ^{-1} we will obtain a continuous map

$$\overline{B}^m(0, 1) \xrightarrow{(\psi, \phi)} \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m).$$

Observe that

$$\|\tilde{a}_{\vec{r}} - c_{\varepsilon \vec{r}}\| = \int_0^1 |\varepsilon a_{\vec{r}}(t\varepsilon) - \varepsilon c_{\vec{r}}(t\varepsilon)| dt = \int_0^\varepsilon |a_{\vec{r}}(t) - c_{\vec{r}}(t)| dt \leq \frac{1}{12} \varepsilon.$$

Using this, the Lipschitz condition for Φ^{-1} (property **(E)**), and the fact that $\Phi^{-1}(c_{\vec{r}}) = (\vec{r}, 0)$ (property **(A)**), we obtain

$$\|\frac{1}{\varepsilon} \psi(\vec{r}) - \vec{r}\| \leq \frac{1}{\varepsilon} \|\psi(\vec{r}) - \varepsilon \vec{r}\| + \frac{1}{\varepsilon} \|\phi(\vec{r}) - 0\| = \frac{1}{\varepsilon} \|\Phi^{-1}(\tilde{a}_{\vec{r}}) - \Phi^{-1}(c_{\varepsilon \vec{r}})\| \leq \frac{6}{\varepsilon} \|\tilde{a}_{\vec{r}} - c_{\varepsilon \vec{r}}\| \leq \frac{1}{2}.$$

In other words, $\tilde{\psi} := \frac{1}{\varepsilon} \psi$ maps a ball $B^m(0, 1)$ continuously into \mathbb{R}^m in such a way that $\|\tilde{\psi}(\vec{r}) - \vec{r}\| \leq \frac{1}{2}$. By Lemma D.7, point $0 \in \mathbb{R}^m$ lies in the image of $\tilde{\psi}$. However, that means that $\psi(\vec{r}_0) = 0$ for some \vec{r}_0 , and hence $\tilde{a}_{\vec{r}_0} = \Phi(0, d)$ for some $d \in \mathcal{BM}_0(I, \mathbb{R}^m)$. By property **(C)**, $[\tilde{a}_{\vec{r}_0}]_{t \in [0, 1]} = 0$. Finally, by Lemma 3.17,

$$0 = [\tilde{a}_{\vec{r}_0}(t)]_{t \in [0, 1]} = [a_{\vec{r}_0}(t)]_{t \in [0, \varepsilon]},$$

which finishes the proof. \square

Second lemma

We will now formulate and prove a result generalising Lemma 8.1. We will work in the same local setting as before. Let us introduce a decomposition $\mathbb{R}^m = \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2}$ and fix $\vec{k}_0 \in \mathbb{R}^m$.

Lemma 8.2. *Let $a_{\vec{r}}(\cdot) \in \mathcal{BM}(I, \mathbb{R}^m)$, where $\vec{r} \in B^{m_1}(0, 1) \oplus \theta_{m_2}$, be a family of E -paths uniformly regular at $t = 0$ w.r.t. \vec{r} and such that $a_{\vec{r}}(0) = \vec{r} + \vec{k}_0$. Let $b_{\vec{s}}(\cdot) \in \mathcal{BM}(I, \mathbb{R}^m)$, where $\vec{s} \in \theta_{m_1} \oplus B^{m_2}(0, 1)$, be a family of E -paths uniformly regular at $t = 0$ w.r.t. \vec{s} and such that $b_{\vec{s}}(0) = \vec{s} + \vec{k}_0$.*

Then there exists a number $\eta > 0$ such that, for every $0 < \varepsilon < \eta$, there exists vectors \vec{r}_0 and \vec{s}_0 ensuring that the curves $a_{\vec{r}_0}(t)$ and $b_{\vec{s}_0}(t)$, after restricting to the interval $[0, \varepsilon]$, realise the same E -homotopy class

$$[a_{\vec{r}_0}(t)]_{t \in [0, \varepsilon]} = [b_{\vec{s}_0}(t)]_{t \in [0, \varepsilon]}.$$

Proof. We will follow the scheme of the final part of the proof of Lemma 8.1, making extensive use of the map $\Phi : \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m) \supset W_0 \rightarrow \Phi(W_0) \subset \mathcal{BM}(I, \mathbb{R}^m)$ constructed before.

By the uniform regularity of the families $a_{\vec{r}}$, $b_{\vec{s}}$, and $c_{\vec{R}}$; there exists a number $\eta > 0$ such that, for $0 < \varepsilon < \eta$,

$$\begin{aligned} \int_0^\varepsilon |a_{\vec{r}}(t) - c_{\vec{r}+\vec{k}_0}(t)| dt &< \frac{1}{24}\varepsilon, & \text{for every } \vec{r} \in B^{m_1}(0,1) \oplus \theta_{m_2} \text{ and} \\ \int_0^\varepsilon |b_{\vec{s}}(t) - c_{\vec{s}+\vec{k}_0}(t)| dt &< \frac{1}{24}\varepsilon, & \text{for every } \vec{s} \in \theta_{m_1} \oplus B^{m_2}(0,1). \end{aligned}$$

Now let us reparametrise the paths $a_{\vec{r}}$ and $b_{\vec{s}}$ by the rule

$$\begin{aligned} \tilde{a}_{\vec{r}}(t) &:= \varepsilon a_{\vec{r}}(\varepsilon t) \quad \text{and} \\ \tilde{b}_{\vec{s}}(t) &:= \varepsilon b_{\vec{s}}(\varepsilon t) \quad \text{for } t \in [0,1]. \end{aligned}$$

We obtained another uniformly regular families of E -paths satisfying $\tilde{a}_{\vec{r}}(0) = \varepsilon(\vec{r} + \vec{k}_0)$ and $\tilde{b}_{\vec{s}}(0) = \varepsilon(\vec{s} + \vec{k}_0)$. Moreover, $\|\tilde{a}_{\vec{r}} - c_{\varepsilon(\vec{r}+\vec{k}_0)}\| \leq \frac{1}{24}\varepsilon$ and $\|\tilde{b}_{\vec{s}} - c_{\varepsilon(\vec{s}+\vec{k}_0)}\| \leq \frac{1}{24}\varepsilon$, and since η can be chosen arbitrary small, we may assume that $\tilde{a}_{\vec{r}}$ and $\tilde{b}_{\vec{s}}$ belong to $\Phi(W_0)$ for all \vec{r} and \vec{s} .

By the uniform regularity of $\tilde{a}_{\vec{r}}$ and $\tilde{b}_{\vec{s}}$, the maps $\tilde{a} : \vec{r} \mapsto \tilde{a}_{\vec{r}}$ and $\tilde{b} : \vec{s} \mapsto \tilde{b}_{\vec{s}}$ are continuous maps from $B^{m_1}(0,1) \oplus \theta_{m_2}$ and $\theta_{m_1} \oplus B^{m_2}(0,1)$, respectively, to $\mathcal{BM}(I, \mathbb{R}^m)$ with L_1 -topology. Composing them with Φ^{-1} we will obtain continuous maps

$$\begin{aligned} B^{m_1}(0,1) \oplus \theta_{m_2} &\xrightarrow{(\psi_a, \phi_a)} \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m) \quad \text{and} \\ \theta_{m_1} \oplus B^{m_2}(0,1) &\xrightarrow{(\psi_b, \phi_b)} \mathbb{R}^m \times \mathcal{BM}_0(I, \mathbb{R}^m). \end{aligned}$$

Now, using the Lipschitz condition for Φ^{-1} and the fact that $\Phi^{-1}(c_{\vec{r}}) = (\vec{r}, 0)$, we can estimate in a way analogous as in the proof of Lemma 8.1 that

$$\begin{aligned} \left\| \frac{1}{\varepsilon} \psi_a(\vec{r}) - (\vec{r} + \vec{k}_0) \right\| &\leq \frac{1}{4} & \text{for } \vec{r} \in B^{m_1}(0,1) \oplus \theta_{m_2} \text{ and} \\ \left\| \frac{1}{\varepsilon} \psi_b(\vec{s}) - (\vec{s} + \vec{k}_0) \right\| &\leq \frac{1}{4} & \text{for } \vec{s} \in \theta_{m_1} \oplus B^{m_2}(0,1). \end{aligned}$$

By Lemma D.8, the images of the maps $\frac{1}{\varepsilon}\psi_a$ and $\frac{1}{\varepsilon}\psi_b$ have a nonempty intersection. In other words, there exist \vec{r}_0 and \vec{s}_0 , vector $\vec{R} \in \mathbb{R}^m$, and $d_a, d_b \in \mathcal{BM}_0(I, \mathbb{R}^m)$ such that $\tilde{a}_{\vec{r}_0} = \Phi(\vec{R}, d_a)$ and $\tilde{b}_{\vec{s}_0} = \Phi(\vec{R}, d_b)$. By property (C), the E -homotopy classes of $\tilde{a}_{\vec{r}_0}$ and $\tilde{b}_{\vec{s}_0}$ are equal. Consequently, by Lemma 3.17,

$$[a_{\vec{r}_0}(t)]_{t \in [0, \varepsilon]} = [\tilde{a}_{\vec{r}_0}(t)]_{t \in [0,1]} = [\tilde{b}_{\vec{s}_0}(t)]_{t \in [0,1]} = [b_{\vec{s}_0}(t)]_{t \in [0, \varepsilon]},$$

which finishes the proof. \square

The proof of the PMP

In this chapter we will finish the proof of Theorems 5.3 and 5.4. In our considerations it is crucial to understand the geometry of the cone \mathbf{K}_τ^u of infinitesimal variations along the optimal trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$. We interpreted \mathbf{K}_τ^u as the set of all directions in $\mathbf{A}_{\mathbf{x}(t_1)}$ in which one can move the point $\mathbf{x}(t_1)$ by performing needle variations of the control u associated with symbols $\mathbf{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$, where τ is fixed. Consequently, a movement in the direction of the ray

$$\mathbf{\Lambda}_{\mathbf{x}(t_1)} := \theta_{\mathbf{x}(t_1)} \oplus \mathbb{R}_+ \cdot (-\partial_t) \subset E_{\mathbf{x}(t_1)} \oplus T_{\underline{\mathbf{x}}(t_1)}\mathbb{R} = \mathbf{A}_{\mathbf{x}(t_1)}$$

would correspond to a variation which decreases the total cost of the trajectory without making changes in the E -evolution. Such a behaviour should not be possible if $\mathbf{f}(\mathbf{x}(t), u(t))$ is a solution of the OCP (\mathbf{P}) , so we may expect that the ray $\mathbf{\Lambda}_{\mathbf{x}(t_1)}$ can be separated from the cone \mathbf{K}_τ^u in such a case. This result is formulated in Theorem 9.1. In the proof we use technical Lemma 8.1 to deduce the existence of E -paths realising certain E -homotopy classes from the infinitesimal picture expressed in the language of the cone \mathbf{K}_τ^u and the ray $\mathbf{\Lambda}_{\mathbf{x}(t_1)}$. When Theorem 9.1 is proved, to finish the proof of Theorem 5.3 we need only to follow a few rather technical steps from the original proof of Pontryagin and his collaborators [Pontryagin *et al.*, 1962].

Theorem 5.4 is proven analogously with some technical modifications. The main difference is that instead of \mathbf{K}_τ^u we use a bigger set \mathbf{K}_τ^u which contains information about both needle variations and initial variations of a given trajectory. Using technical Lemma 8.2 we prove Theorem 9.6 describing the geometry of \mathbf{K}_τ^u . Then, basing on this result, we make a few final steps after [Pontryagin *et al.*, 1962].

9.1 The proof of Theorem 5.3

The geometry of the cone \mathbf{K}_τ^u

Throughout this section we assume that the controlled pair $(\mathbf{x}(t), u(t))$ is a solution of the OCP (\mathbf{P}) . All results obtained in this section are valid under this assumption.

Theorem 9.1. *Let $(\mathbf{x}(t), u(t))$, for $t \in [t_0, t_1]$, be a solution of the optimal control problem (\mathbf{P}) . Then the ray $\mathbf{\Lambda}_{\mathbf{x}(t_1)}$ and the convex cone \mathbf{K}_τ^u can be separated for any $\tau \in (t_0, t_1)$, which is a regular point of u .*

The idea of the proof is the following. Assuming the contrary we will construct a family of symbols $\mathfrak{w}(\vec{r})$, where $\vec{r} \in E_{x(t_1)}$, such that the associated infinitesimal variations $\mathbf{d}^{\mathfrak{w}(\vec{r})}(s)$ are uniformly regular w.r.t. \vec{r} at $s = 0$ and point into the directions $\vec{r} - \partial_t \in E_{x(t_1)} \oplus T_{\underline{x}(t_1)}\mathbb{R} = \mathbf{A}_{x(t_1)}$. For such a family we will be able to use Lemma 8.1 to deduce that for some \vec{r}_0 the variation $\mathbf{d}^{\mathfrak{w}(\vec{r}_0)}(s)$ has special properties. Next we will show that in such a case the pair $(\mathbf{x}(t), u(t))$ cannot be a solution of the OCP (P).

Proof. Assume the contrary, i.e., that the convex cone \mathbf{K}_τ^u and the ray $\mathbf{A}_{x(t_1)}$ cannot be separated. Denote by $\boldsymbol{\lambda} := \theta_{x(t_1)} - \partial_t \in E_{x(t_1)} \oplus T_{\underline{x}(t_1)}\mathbb{R}$ a vector spanning $\mathbf{A}_{x(t_1)}$. It follows from Lemma D.6 that there exists a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ of $E_{x(t_1)} \oplus T_{\underline{x}(t_1)} \subset \mathbf{A}_{x(t_1)}$ such that vectors $\boldsymbol{\lambda}, \boldsymbol{\lambda} + \mathbf{e}_i, \boldsymbol{\lambda} - \mathbf{e}_i$ lie in \mathbf{K}_τ^u for $i = 1, \dots, m$.

Denote by \mathfrak{w} , \mathfrak{w}_i , and \mathfrak{v}_i some symbols such that elements in \mathbf{K}_τ^u corresponding to these symbols are $\mathbf{d}^{\mathfrak{w}}(0) = \boldsymbol{\lambda}$, $\mathbf{d}^{\mathfrak{w}_i}(0) = \boldsymbol{\lambda} + \mathbf{e}_i$, and $\mathbf{d}^{\mathfrak{v}_i}(0) = \boldsymbol{\lambda} - \mathbf{e}_i$, respectively. We deal with a finite set of symbols, hence we can assume that they all are of the form $(\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$, where τ_i, v_i, k , and τ are fixed, and that they differ by δt_i and δt (we can always add a triple $(\tau_i, v_i, \delta t_i = 0)$ to a symbol without changing anything — cf. Remark 7.3). For any $\vec{r} = \sum_{i=1}^m r^i \mathbf{e}_i \in B^m(0, 1) = \{\vec{r} : \|\vec{r}\| = \sum_i |r_i| \leq 1\}$ we may define a new symbol

$$\mathfrak{w}(\vec{r}) = \left(1 - \sum_{i=1}^m |r^i|\right) \mathfrak{w} + \sum_{i=1}^m h^+(r^i) \mathfrak{w}_i + \sum_{i=1}^m h^-(r^i) \mathfrak{v}_i,$$

where $h^+(r) = \max\{r, 0\}$ and $h^-(r) = \max\{-r, 0\}$ are non-negative, and the convex combination of symbols is defined using the natural rule

$$\begin{aligned} & \nu(\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k} + \mu(\tau_i, v_i, \tau, \delta t'_i, \delta t')_{i=1, \dots, k} \\ &= (\tau_i, v_i, \tau, \nu \delta t_i + \mu \delta t'_i, \nu \delta t + \mu \delta t')_{i=1, \dots, k}. \end{aligned}$$

We will now study the properties of \mathbf{A} -paths $s \mapsto \mathbf{d}^{\mathfrak{w}(\vec{r})}(s)$ corresponding to symbols $\mathfrak{w}(\vec{r})$ (see Remark 7.7).

From (7.3) it is straightforward to verify that

$$\mathbf{d}^{\mathfrak{w}(\vec{r})}(0) = \boldsymbol{\lambda} + \vec{r}.$$

It follows from Theorem 7.4 that, since the numbers $\delta t(\vec{r})$ and $\delta t_i(\vec{r})$ in the symbol $\mathfrak{w}(\vec{r})$ depend continuously on \vec{r} , which takes values in a compact set, we may choose $\theta > 0$ such that $\mathbf{d}^{\mathfrak{w}(\vec{r})}(s)$ is well-defined for $s \in [0, \theta]$ and all $\vec{r} \in B^m(0, 1)$,

Now consider $\underline{\mathbf{d}}^{\mathfrak{w}(\vec{r})}(s)$ — the projections of the family of \mathbf{A} -paths $\mathbf{d}^{\mathfrak{w}(\vec{r})}(s)$ to the algebroid $T\mathbb{R}$. Observe that since $\mathbf{d}^{\mathfrak{w}(\vec{r})}(s)$ are uniformly regular, so are $\underline{\mathbf{d}}^{\mathfrak{w}(\vec{r})}(s)$. Since in canonical coordinates on $T\mathbb{R}$ we have $\underline{\mathbf{d}}^{\mathfrak{w}(\vec{r})}(0) = -1$, there exist a number $0 < \eta \leq \theta$ such that

$$\int_0^\varepsilon \underline{\mathbf{d}}^{\mathfrak{w}(\vec{r})}(s) ds < 0 \tag{9.1}$$

for all $\varepsilon \leq \eta$ and all $\vec{r} \in B^m(0, 1)$. This property will be used later.

After projecting $\underline{d}^{\mathfrak{w}(\vec{r})}(s)$ from \mathbf{A} onto E , we obtain a family of bounded measurable admissible paths $\underline{d}^{\mathfrak{w}(\vec{r})}(s)$, again uniformly regular at $s = 0$ w.r.t. $\vec{r} \in B^m(0, 1) \subset \mathbb{R}^m$, and such that $\underline{d}^{\mathfrak{w}(\vec{r})}(0) = \vec{r} \in \mathbb{R}^m \approx E_{x(t_1)}$. In other words, the paths $\underline{d}^{\mathfrak{w}(\vec{r})}(\cdot)$ satisfy the assumptions of Lemma 8.1 and, consequently, there exists a vector $\vec{r}_0 \in B^m(0, 1)$ and a number $0 < \varepsilon < \eta$ such that

$$[\underline{d}^{\mathfrak{w}(\vec{r}_0)}(t)]_{t \in [0, \varepsilon]} = [0]. \quad (9.2)$$

Properties (9.1) and (9.2) contradict the optimality of $\mathbf{f}(\mathbf{x}(t), u(t))$. Indeed, from (7.14) we know that $\mathbf{f}(\mathbf{x}(t, s), u_s(t))$ — the variation of the trajectory $\mathbf{f}(\mathbf{x}(t), u(t))$ associated with a needle variation $u_s(t) := u_s^{\mathfrak{w}(\vec{r}_0)}(t)$ satisfies

$$[\mathbf{f}(\mathbf{x}(t, \varepsilon), u_\varepsilon(t))]_{t \in [t_0, t_1 + \varepsilon \delta t(\vec{r}_0)]} = [\mathbf{f}(\mathbf{x}(t), u(t))]_{t \in [t_0, t_1]} [\underline{d}^{\mathfrak{w}(\vec{r}_0)}(s)]_{s \in [0, \varepsilon]}.$$

Projecting the above equality to the algebroid E and using (9.2) we get

$$[f(x(t, \varepsilon), u_\varepsilon(t))]_{t \in [t_0, t_1 + \varepsilon \delta t(\vec{r}_0)]} = [f(x(t), u(t))]_{t \in [t_0, t_1]},$$

and hence the E -homotopy classes agree.

What is more, the \mathbb{R} -projection gives

$$[L(x(t, \varepsilon), u_\varepsilon(t))]_{t \in [t_0, t_1 + \varepsilon \delta t(\vec{r}_0)]} = [L(x(t), u(t))]_{t \in [t_0, t_1]} \left[\underline{d}^{\mathfrak{w}(\vec{r}_0)}(s) \right]_{s \in [0, \varepsilon]}.$$

From (9.1) we deduce that the total costs satisfy the following inequality:

$$\begin{aligned} \underline{x}(t_1 + \varepsilon \delta t(\vec{r}_0), \varepsilon) &= \int_{t_0}^{t_1 + \varepsilon \delta t(\vec{r}_0)} L(x(s, \varepsilon), u_\varepsilon(s)) ds = \int_{t_0}^{t_1} L(x(s), u(s)) ds + \int_0^\varepsilon \underline{d}^{\mathfrak{w}(\vec{r}_0)}(s) ds \\ &\stackrel{(9.1)}{<} \int_{t_0}^{t_1} L(x(s), u(s)) ds = \underline{x}(t_1). \end{aligned}$$

The above inequality proves that $\mathbf{f}(\mathbf{x}(t), u(t))$ cannot be a solution of the OCP (\mathbf{P}) , which stays in a contradiction to our assumptions. \square

To finish the proof of Theorem 5.3 we will now follow the steps of the original result of [Pontryagin *et al.*, 1962]. All the important information is contained in Theorem 7.4 telling us that the set of infinitesimal variations \mathbf{K}_τ^u is a convex cone with elements defined by means of a local one-parameter group $\mathbf{B}_{t\tau_0}$ (see (7.3)) and in Theorem 9.1 describing the geometry of this cone. The structure of an AL algebroid, necessary to prove the above results, will now play no essential role.

The construction of $\xi(t)$ and the “maximum principle”

Fix an element $\tau \in (t_0, t_1)$ to be a regular point of u . In view of Theorem 9.1 there exists a non-zero covector $\xi(t_1) \in \mathbf{A}_{x(t_1)}^*$ separating \mathbf{K}_τ^u and $\mathbf{A}_{x(t_1)}$; that is (confront Remark D.5),

$$\langle \underline{d}, \xi(t_1) \rangle_\tau \leq 0 \leq \langle \lambda, \xi(t_1) \rangle_\tau \quad \text{for every } \underline{d} \in \mathbf{K}_\tau^u. \quad (9.3)$$

Let us define $\xi(t) := \mathbf{B}_{t\tau_1}^*(\xi(t_1)) \in \mathbf{A}_{x(t)}^*$ for $t \in [t_0, t_1]$.

Lemma 9.2. *For every $t \in [t_0, \tau]$ which is a regular point of the control $u(t)$ the following “maximum principle” holds:*

$$\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)) = \sup_{v \in U} \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v).$$

Moreover, $\mathbf{H}(\mathbf{x}(\tau), \boldsymbol{\xi}(\tau), u(\tau)) = 0$.

Proof. Choose a regular point $t \in [t_0, \tau]$, take an arbitrary element $v \in U$ and a number $\delta t_1 > 0$, and consider a symbol $\mathbf{w} = (\tau_1 = t, \delta t_1, v_1 = v, \tau, \delta t = 0)$. The corresponding element $\mathbf{d}^{\mathbf{w}}(0) \in \mathbf{K}_\tau^u$ equals $\mathbf{B}_{t_1 t}[\mathbf{f}(\mathbf{x}(t), v) - \mathbf{f}(\mathbf{x}(t), u(t))]\delta t_1$ (cf. (7.3)). From (9.3) we obtain

$$\begin{aligned} 0 &\geq \langle \mathbf{B}_{t_1 t}[\mathbf{f}(\mathbf{x}(t), v) - \mathbf{f}(\mathbf{x}(t), u(t))], \boldsymbol{\xi}(t_1) \rangle_\tau \delta t_1 \\ &\stackrel{\text{rem. 4.6}}{=} \langle \mathbf{B}_{t t_1} \mathbf{B}_{t_1 t}[\mathbf{f}(\mathbf{x}(t), v) - \mathbf{f}(\mathbf{x}(t), u(t))], \mathbf{B}_{t t_1}^*(\boldsymbol{\xi}(t_1)) \rangle_\tau \delta t_1 \\ &= \langle \mathbf{f}(\mathbf{x}(t), v) - \mathbf{f}(\mathbf{x}(t), u(t)), \boldsymbol{\xi}(t) \rangle_\tau \delta t_1 \\ &= \left(\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v) - \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)) \right) \delta t_1. \end{aligned}$$

Since $\delta t_1 > 0$, we have $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v) \leq \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t))$ for arbitrarily chosen $v \in U$.

To prove the second part of the assertion, consider a symbol $\mathbf{v} = (\tau, \delta t)$. The associated element $\mathbf{d}^{\mathbf{v}}(0)$ is $\mathbf{B}_{t_1 \tau}(\mathbf{f}(\mathbf{x}(\tau), u(\tau)))\delta t \in \mathbf{K}_\tau^u$. Consequently, from (9.3), we obtain

$$\begin{aligned} 0 &\geq \langle \mathbf{B}_{t_1 \tau}(\mathbf{f}(\mathbf{x}(\tau), u(\tau))), \boldsymbol{\xi}(t_1) \rangle_\tau \delta t \\ &\stackrel{\text{rem. 4.6}}{=} \langle \mathbf{f}(\mathbf{x}(\tau), u(\tau)), \boldsymbol{\xi}(\tau) \rangle_\tau \delta t = \mathbf{H}(\mathbf{x}(\tau), \boldsymbol{\xi}(\tau), u(\tau))\delta t. \end{aligned}$$

Since δt can be arbitrary, we deduce that $\mathbf{H}(\mathbf{x}(\tau), \boldsymbol{\xi}(\tau), u(\tau)) = 0$. □

The condition $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)) = 0$

To finish the proof just two more things are left. We have to check that the Hamiltonian $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t))$ is constantly 0 along the optimal trajectory, and we have to extend the “maximum principle” to all regular $t \in [t_0, t_1]$ (so far it holds only on the interval $[t_0, \tau]$, where $\tau < t_1$ is a fixed regular point).

Lemma 9.3. *For $\boldsymbol{\xi}(t)$ defined as above, the equality $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)) = 0$ holds at every regular point $t \in [t_0, \tau]$ of the control u .*

Proof. Denote by P the closure of the set $\{u(t) : t \in [t_0, \tau]\}$. Since $u(t)$ is bounded, P is a compact subset of U . Define a new function $\mathbf{m} : \mathbf{A}^* \rightarrow \mathbb{R}$ by the formula

$$\mathbf{m}(\mathbf{x}, \boldsymbol{\xi}) := \max_{v \in P} \mathbf{H}(\mathbf{x}, \boldsymbol{\xi}, v).$$

It follows from the previous lemma that $\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t)) = \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t))$ at every regular point t of u . We shall show that $\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t))$ is constant on $[t_0, \tau]$, and hence equals $\mathbf{H}(\mathbf{x}(\tau), \boldsymbol{\xi}(\tau), u(\tau)) = 0$ (confront Lemma 9.2). Observe that the function $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v)$

is uniformly (for all $v \in P$) Lipschitz w.r.t. t . Indeed, in local coordinates $(x^a, \underline{x}, y^i, \underline{y})$ on \mathbf{A} and $(x^a, \underline{x}, \xi_i, \underline{\xi})$ on \mathbf{A}^* we have $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v) = \xi_i(t) f^i(x(t), v) + \underline{\xi} L(x(t), v)$. Note that, by assumption, functions $f^i(x, v)$ and $L(x, v)$ are C^1 w.r.t. x and their x -derivatives are continuous functions of both variables. Since $x(t)$ is an AC path with bounded derivative, functions $\frac{\partial f^i}{\partial x^a}(x(t), v)$ and $\frac{\partial L}{\partial x^a}(x(t), v)$, as well as functions $f^i(x(t), v)$ and $L(x(t), v)$, are bounded in $[t_0, \tau] \times P$. As the evolution of $\xi(t)$ is governed by (4.7) and $u(t)$ is bounded on $[t_0, \tau]$, the derivatives $\partial_t \xi^k(t)$ are also bounded on $[t_0, \tau]$. Consequently, since the path $(\mathbf{x}(t), \boldsymbol{\xi}(t)) \in \mathbf{A}^*$ can be covered by a finite number of coordinate charts, the t -derivative of $\mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v)$ is bounded on $[t_0, \tau] \times P$. As a result there exists a number C such that

$$\left| \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v) - \mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), v) \right| \leq C|t - t'|$$

for all $t, t' \in [t_0, \tau]$ and for any $v \in P$. Observe also that

$$\begin{aligned} \frac{\partial}{\partial t} \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), v) |_{v=u(t)} &= \partial_t \xi_i(t) f^i(x(t), u(t)) + \\ &+ \xi_i(t) \partial_t f^i(x(t), v) |_{v=u(t)} + \underline{\xi} \partial_t L(x(t), v) |_{v=u(t)} \stackrel{(4.7)}{=} \\ &= \left[-\rho_i^a(x) \left(\frac{\partial f^k}{\partial x^a}(x, u(t)) \xi_k(t) + \frac{\partial L}{\partial x^a}(x, u(t)) \underline{\xi}(t) \right) + c_{ji}^k(x) f^j(x, u(t)) \xi_k(t) \right] f^i(x, u(t)) \\ &+ \xi_i(t) \frac{\partial f^i}{\partial x^a}(x(t), u(t)) \rho_k^a(x) f^k(x(t), u(t)) + \underline{\xi} \frac{\partial L}{\partial x^a}(x(t), u(t)) \rho_k^a(x) f^k(x(t), u(t)) \\ &= \xi_k(t) c_{ij}^k(x) f^i(x(t), u(t)) f^j(x(t), u(t)) = 0 \end{aligned}$$

by the skew-symmetry of $c_{ij}^k(x)$.

Now take any regular points $t, t' \in [t_0, \tau]$. Since $u(t), u(t') \in P$, we have

$$\begin{aligned} -C|t - t'| &\leq \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t')) - \mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), u(t')) \\ &\leq \mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t)) - \mathbf{m}(\mathbf{x}(t'), \boldsymbol{\xi}(t')) \\ &= \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)) - \mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), u(t')) \\ &\leq \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)) - \mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), u(t)) \leq C|t - t'|; \end{aligned}$$

i.e., $\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t))$ satisfies the Lipschitz condition on the set of regular points (dense in $[t_0, \tau]$). It is also a continuous map (since $x(t)$ and $\xi(t)$ are AC and the maps $L(x, v)$, $f(x, v)$ are continuous in both variables), therefore it is Lipschitz on the whole interval $[t_0, \tau]$. By Rademacher's theorem, $\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t))$ is almost everywhere differentiable on $[t_0, \tau]$. Now take any point t of differentiability of $\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t))$ which is also a point of the regularity of the control u . We have

$$\mathbf{m}(\mathbf{x}(t'), \boldsymbol{\xi}(t')) - \mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t)) \geq \mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), u(t)) - \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t)).$$

For $t' > t$, we get

$$\frac{\mathbf{m}(\mathbf{x}(t'), \boldsymbol{\xi}(t')) - \mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t))}{t' - t} \geq \frac{\mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), u(t)) - \mathbf{H}(\mathbf{x}(t), \boldsymbol{\xi}(t), u(t))}{t' - t}.$$

Consequently,

$$\frac{d}{dt}\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t)) \geq \frac{\partial}{\partial t'} \Big|_{t'=t} \mathbf{H}(\mathbf{x}(t'), \boldsymbol{\xi}(t'), u(t)) = 0.$$

Similarly, for $t' < t$, we would get $\frac{d}{dt}\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t)) \leq 0$. We deduce that $\frac{d}{dt}\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t)) = 0$ a.e. in $[t_0, \tau]$; hence $\mathbf{m}(\mathbf{x}(t), \boldsymbol{\xi}(t))$ is constant and equals $\mathbf{H}(\mathbf{x}(\tau), \boldsymbol{\xi}(\tau), u(\tau)) = 0$. \square

Extending the "maximum principle" to $[t_0, t_1]$

Lemma 9.4. *Let $t \in [t_0, \tau]$ be any regular point of the control u . Then*

$$\mathbf{K}_t^u \subset \text{cl}(\mathbf{K}_\tau^u).$$

Proof. Consider an element $\mathbf{d} \in \mathbf{K}_t^u$ of the form

$$\mathbf{d} = \mathbf{B}_{t_1 t}[\mathbf{f}(\mathbf{x}(t), u(t))]\delta t + \sum_{i=1}^s \mathbf{B}_{t_1 \tau_i}[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i))]\delta t_i.$$

Since $\text{cl}(\mathbf{K}_\tau^u)$ is a convex cone it is enough to show that $\mathbf{B}_{t_1 t}(\mathbf{f}(\mathbf{x}(t), u(t))\delta t)$ and $\sum_{i=1}^s \mathbf{B}_{t_1 \tau_i}[\mathbf{f}(\mathbf{x}(\tau_i), v_i) - \mathbf{f}(\mathbf{x}(\tau_i), u(\tau_i))]\delta t_i$ belong to $\text{cl}(\mathbf{K}_\tau^u)$. The later clearly belongs to $\mathbf{K}_\tau^u \subset \text{cl}(\mathbf{K}_\tau^u)$ since $\tau_i \leq t < \tau$.

Assume now that $\mathbf{B}_{t_1 t}[\mathbf{f}(\mathbf{x}(t), u(t))\delta t]$ does not belong to $\text{cl}(\mathbf{K}_\tau^u)$. Since this set is a closed convex cone, by Theorem D.2, there exists a covector $\bar{\boldsymbol{\xi}}(t_1) \in \mathbf{A}_{\mathbf{x}(t_1)}^*$ strictly separating $\text{cl}(\mathbf{K}_\tau^u)$ from $\{\mathbf{B}_{t_1 t}(\mathbf{f}(\mathbf{x}(t), u(t))\delta t)\}$; i.e.,

$$\langle \mathbf{d}, \bar{\boldsymbol{\xi}}(t_1) \rangle_\tau \leq 0 < \langle \mathbf{B}_{t_1 t}(\mathbf{f}(\mathbf{x}(t), u(t))\delta t), \bar{\boldsymbol{\xi}}(t_1) \rangle_\tau \quad \text{for any } \mathbf{d} \in \text{cl}(\mathbf{K}_\tau^u).$$

Define $\bar{\boldsymbol{\xi}}(t) := \mathbf{B}_{t t_1}^*(\bar{\boldsymbol{\xi}}(t_1))$. Lemmas 9.2 and 9.3 hold for $\bar{\boldsymbol{\xi}}(t)$ (we needed only $\langle \mathbf{d}, \bar{\boldsymbol{\xi}}(t_1) \rangle \leq 0$ for $\mathbf{d} \in \mathbf{K}_\tau^u \subset \text{cl}(\mathbf{K}_\tau^u)$ in the proofs), hence, in particular,

$$\langle \mathbf{f}(\mathbf{x}(t), u(t)), \bar{\boldsymbol{\xi}}(t) \rangle_\tau = \mathbf{H}(\mathbf{x}(t), \bar{\boldsymbol{\xi}}(t), u(t)) = 0,$$

as t is a regular point of u . On the other hand,

$$0 < \langle \mathbf{B}_{t_1 t}(\mathbf{f}(\mathbf{x}(t), u(t))\delta t), \bar{\boldsymbol{\xi}}(t_1) \rangle_\tau = \langle \mathbf{f}(\mathbf{x}(t), u(t)), \bar{\boldsymbol{\xi}}(t) \rangle_\tau \delta t = \mathbf{H}(\mathbf{x}(t), \bar{\boldsymbol{\xi}}(t), u(t))\delta t,$$

and hence $\mathbf{H}(\mathbf{x}(t), \bar{\boldsymbol{\xi}}(t), u(t)) \neq 0$. This contradiction finishes the proof. \square

Note that, so far, we could define the set \mathbf{K}_τ^u only for a regular point $\tau < t_1$. With the help of the above lemma we can also define $\mathbf{K}_{t_1}^u$ as the direct limit of the increasing family of sets $\text{cl}(\mathbf{K}_\tau^u)$,

$$\mathbf{K}_{t_1}^u := \bigcup_{\tau < t_1, \tau \text{ regular}} \text{cl}(\mathbf{K}_\tau^u).$$

It is clear that $\mathbf{K}_{t_1}^u$ is a convex cone in $\mathbf{A}_{\mathbf{x}(t_1)}$. It has geometric properties analogous to the properties of \mathbf{K}_τ^u described in Theorem 9.1.

Lemma 9.5. *The ray $\Lambda_{\mathbf{x}(t_1)}$ and the convex cone $\mathbf{K}_{t_1}^u$ are separable.*

Proof. Assume the contrary. By Lemma D.6 (we use it for $V = \mathbf{A}_{\mathbf{x}(t_1)} = W \oplus \mathbb{R} = E_{\mathbf{x}(t_1)} \oplus T_{\underline{\mathbf{x}}(t_1)}\mathbb{R}$, $K_1 = \mathbf{K}_{t_1}^u$, $S = \{0\} \subset W$, and $K_2 = \Lambda_{\mathbf{x}(t_1)}$), there exists a vector $k \in \mathbf{K}_{t_1}^u \cap \Lambda_{\mathbf{x}(t_1)}$ and vectors $e_1, \dots, e_m \in E_{\mathbf{x}(t_1)}$ such that (D.3) and (D.4) hold. Since $\mathbf{K}_{t_1}^u$ is a limit of an increasing family of sets, there exists a regular $t < t_1$ such that (D.3) and (D.4) hold for $K_1 = \text{cl}(\mathbf{K}_t^u)$. In other words, $\Lambda_{\mathbf{x}(t_1)}$ and $\text{cl}(\mathbf{K}_t^u)$ are not separable. By Lemma D.3 also $\Lambda_{\mathbf{x}(t_1)}$ and \mathbf{K}_t^u are not separable. This contradicts Theorem 9.1. \square

Now choose a non-zero covector $\xi(t_1) \in \mathbf{A}_{\mathbf{x}(t_1)}^*$ separating $\mathbf{K}_{t_1}^u$ and $\Lambda_{\mathbf{x}(t_1)}$ and define $\xi(t) := \mathbf{B}_{tt_1}^*(\xi(t_1))$. We have

$$\langle d, \xi(t_1) \rangle_\tau \leq 0 \leq \langle \lambda, \xi(t_1) \rangle_\tau \quad \text{for every } d \in \mathbf{K}_{t_1}^u.$$

Since, by construction, $\mathbf{K}_\tau^u \subset \mathbf{K}_{t_1}^u$, the covector $\xi(t_1)$ separates also \mathbf{K}_τ^u and $\Lambda_{\mathbf{x}(t_1)}$. This is enough for Lemmas 9.2 and 9.3 to hold for $\xi(t)$. As a consequence, for every regular point of the control u , we have

$$\mathbf{H}(\mathbf{x}(t), \xi(t), u(t)) = \sup_{v \in U} \mathbf{H}(\mathbf{x}(t), \xi(t), v) = 0.$$

Finally, since $\langle \lambda, \xi(t_1) \rangle \geq 0$, we have $\underline{\xi}(t) \equiv \underline{\xi}(t_1) \leq 0$. This finishes the proof of Theorem 5.3. \square

9.2 The proof of Theorem 5.4

In this section we assume that the controlled pair $(\mathbf{x}(t), u(t))$, with $t \in [t_0, t_1]$, solves the OCP (**P rel**). Recall that we assume that $\mathbf{x}(t_0) = (x(t_0), \underline{x}(t_0)) = (\phi_0(z_0), 0)$ and $\mathbf{x}(t_1) = (x(t_1), \underline{x}(t_1)) = (\phi_1(w_0), \underline{x}(t_1))$, where ϕ_i are base projections of algebroid morphisms $\Phi_i : TS_i \rightarrow E$ for $i = 0, 1$. We use the following notation $\mathcal{S}_0 := \Phi_0(T_{z_0}S_0)$, $\mathcal{S}_1 := \Phi_1(T_{w_0}S_1)$, $\Lambda_{\underline{x}} := \mathbb{R}_+(-\partial_t) \subset T_{\underline{x}}\mathbb{R}$, and $\mathcal{K}_\tau^u := \text{conv} \{ \mathbf{B}_{t_1 t_0}(\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)}), \mathbf{K}_\tau^u \}$.

Theorem 9.6. *Let $(\mathbf{x}(t), u(t))$, for $t \in [t_0, t_1]$, be a solution of the optimal control problem (**P rel**). Then the convex cones \mathcal{K}_τ^u and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ can be separated for any $\tau \in (t_0, t_1)$ which is a regular point of u .*

Proof. The argument is very similar to that from Theorem 9.1.

Assume that $(\mathbf{x}(t), u(t))$ is a solution of (**P rel**) but the cones \mathcal{K}_τ^u and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ are not separable. First, construct a family of admissible paths $s \mapsto b_0(s, p) \in E$ parametrised by $p \in \mathcal{S}_0$ such that the following conditions hold:

- the family is uniformly regular w.r.t p at $s = 0$,
- $b_0(0, p) = p$,
- $b_0(s, p)$ lies in the image of Φ_0 .

A family with the desired properties can be build as follows. First, choose a linear subspace $V \subset T_{z_0}S_0$ such that $\Phi_0|_V : V \rightarrow \mathcal{S}_0 = \Phi_0(T_{z_0}S_0)$ is a linear isomorphism (Φ_0 at z_0 is a linear map from $T_{z_0}S_0$ to $\mathcal{S}_0 \subset E_{x(t_0)}$). For $p \in \mathcal{S}_0$ take $v(p) = (\Phi_0|_V)^{-1}(p) \in V$ and consider a curve $s \mapsto z(s, p) = \exp_{z_0}(s \cdot v(p))$ in S_0 , where $\exp_{z_0}(\cdot)$ is an exponential map around z_0 defined for some metric on S_0 . We define $b_0(s, p) := \Phi_0(\partial_s z(s, p))$.

Let us check that $b_0(s, p)$ has the desired properties. Clearly, it lies in the image of Φ_0 . The properties of the exponential map imply that $s \mapsto \tilde{z}(s, v) = \partial_s(\exp_{z_0}(s \cdot v))$ is a family of paths uniformly regular at $s = 0$ w.r.t v . These paths are admissible in the tangent algebroid TS_0 . Now, since $b_0(s, p)$ is obtained as a composition of $\tilde{z}(s, v)$ with a continuous map $p \mapsto v(p)$ and a smooth map Φ_0 , the uniform regularity is preserved. Since Φ_0 is an algebroid morphism, also admissibility is preserved. Finally, observe that $b_0(0, p) = \Phi_0[\partial_s|_0 \exp_{z_0}(s \cdot v(p))] = \Phi_0(v(p)) = \Phi_0((\Phi_0|_V)^{-1})(p) = p$.

Similarly, we construct a family of admissible paths $s \mapsto b_1(s, q) \in E$ parametrised by $q \in \mathcal{S}_1$ and such that the following holds:

- the family is uniformly regular w.r.t q at $s = 0$,
- $b_1(0, q) = q$,
- $b_1(s, q)$ lies in the image of Φ_1 .

If the cones \mathcal{K}_τ^u and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ are not separable, by Lemma D.6, there exists a vector $\mathbf{k} \in \mathcal{K}_\tau^u \cap (\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)})$ and vectors $e_1, \dots, e_{m_1} \in E_{x(t_1)} \oplus \theta_{\underline{x}(t_1)}$ such that

- $\text{span}\{e_1, \dots, e_{m_1}, \mathcal{S}_1\} = E_{x(t_1)}$,
- $\mathbf{k} \pm e_i \in \mathcal{K}_\tau^u$ for $i = 1, \dots, m_1$.

In the last formula we understand $\mathbf{k} + e_i$ as $\mathbf{k} + (e_i + \theta_{\underline{x}(t_1)})$, where $e_i + \theta_{\underline{x}(t_1)} \in E_{x(t_1)} \oplus T_{\underline{x}(t_1)} = \mathbf{A}_{x(t_1)}$. In view the first property, we can decompose $E_{x(t_1)} \approx \mathbb{R}^m = \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2} = \text{span}\{e_1, \dots, e_{m_1}\} \oplus \mathcal{S}_1$ and choose a basis $(\tilde{e}_1, \dots, \tilde{e}_{m_2})$ of \mathcal{S}_1 .

Without the loss of generality (cones are invariant under rescaling) we may assume that \mathbf{k} projects to $-\partial_t$ under $p_{\mathbb{T}\mathbb{R}} : \mathbf{A} = E \times \mathbb{T}\mathbb{R} \rightarrow \mathbb{T}\mathbb{R}$. Observe that the E -projection $\vec{k}_0 := p_E(\mathbf{k})$ belongs to \mathcal{S}_1 .

Introduce $\mathbf{b}_0(s, p) := (b_0(s, p), 0) \in E \times \mathbb{T}\mathbb{R} = \mathbf{A}$. We can consider variations $\mathbf{f}(\mathbf{x}(t, s, p), u_s^\mathfrak{w}(t))$ associated with symbols $\mathfrak{w} = (\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$ and initial base-point variations $\mathbf{x}_0(s, p) = \boldsymbol{\tau}(\mathbf{b}_0(s, p))$ as in Theorem 7.4. From (7.2), there exists a family of \mathbf{A} -paths $s \mapsto \mathbf{d}^{\mathfrak{w}, p}(s)$ defined for $0 \leq s \leq \theta$ such that

$$\begin{aligned} & [\mathbf{b}_0(s, p)]_{s \in [0, \varepsilon]} [\mathbf{f}(\mathbf{x}(t, \varepsilon, p)), u_\varepsilon^\mathfrak{w}(t)]_{t \in [t_0, t_1 + \varepsilon \delta t]} \\ &= [\mathbf{f}(\mathbf{x}(t), u(t))]_{t \in [t_0, t_1]} [\mathbf{d}^{\mathfrak{w}, p}(s)]_{s \in [0, \varepsilon]}, \end{aligned} \tag{9.4}$$

Observe that, due to (7.3), $\mathbf{d}^{\mathfrak{w}, p}(0) \in \mathcal{K}_\tau^u$ and, moreover, all elements of \mathcal{K}_τ^u can be obtained in this way.

Choose symbols \mathfrak{w} , \mathfrak{w}_i , and \mathfrak{v}_i and elements $p, p_i, \tilde{p}_i \in \mathcal{S}_0$ such that $\mathbf{d}^{\mathfrak{w}, p}(0) = \mathbf{k}$, $\mathbf{d}^{\mathfrak{w}_i, p_i}(0) = \mathbf{k} + \mathbf{e}_i$, and $\mathbf{d}^{\mathfrak{v}_i, \tilde{p}_i}(0) = \mathbf{k} - \mathbf{e}_i$. Since we deal with a finite set of symbols, we can assume that they all are of the form $(\tau_i, v_i, \tau, \delta t_i, \delta t)_{i=1, \dots, k}$, where τ_i , v_i , τ , and k are

fixed, and that they differ by δt_i and δt (see Remark 7.3). Define for any $\vec{r} = \sum_{i=1}^{m_1} r^i e_i \in B^{m_1}(0, 1) \oplus \theta_{m_2} = \{\vec{r} : \|\vec{r}\| = \sum_i |r_i| \leq 1\} \subset E_{x(t_1)}$ a new symbol

$$\mathfrak{w}(\vec{r}) = \left(1 - \sum_{i=1}^m |r^i|\right) \mathfrak{w} + \sum_{i=1}^m h^+(r^i) \mathfrak{w}_i + \sum_{i=1}^m h^-(r^i) \mathfrak{v}_i,$$

and a new element of \mathcal{S}_1

$$p(\vec{r}) = \left(1 - \sum_{i=1}^{m_1} |r^i|\right) p + \sum_{i=1}^{m_1} h^+(r^i) p_i + \sum_{i=1}^{m_1} h^-(r^i) \tilde{p}_i,$$

where $h^+(r) = \max\{r, 0\}$ and $h^-(r) = \max\{-r, 0\}$ are non-negative. From (7.3) we get

$$\mathbf{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(0) = \mathbf{k} + \vec{r}.$$

Now from Theorem 7.4 there exists $\theta > 0$ such that $\mathbf{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(s)$ is well-defined for $0 \leq s \leq \theta$ and all $\vec{r} \in B^{m_1}(0, 1) \oplus \theta_{m_2}$ and, moreover, it is uniformly regular w.r.t. \vec{r} at $s = 0$.

Repeating the reasoning from the proof of Theorem 9.1 we can show that the $\mathbb{T}\mathbb{R}$ -projections $\underline{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(s)$ satisfy

$$\int_0^\varepsilon \underline{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(s) ds < 0 \quad (9.5)$$

for all $\varepsilon \leq \eta$ and all $\vec{r} \in B^{m_1}(0, 1)$, where $0 < \eta \leq \theta$ is a fixed number.

Projecting $\mathbf{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(s)$ to E we obtain a family of bounded measurable admissible paths $\mathbf{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(s)$ uniformly regular at $s = 0$ w.r.t. $\vec{r} \in B^{m_1}(0, 1) \oplus \theta_{m_2} \subset \mathbb{R}^m$, and such that $\mathbf{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(0) = \vec{r} + \vec{k}_0 \in \mathbb{R}^m \approx E_{x(\tau)}$.

For

$$\vec{s} \in \theta_{m_1} \oplus B^{m_2}(0, 1) = \{\vec{s} = \sum_{i=1}^{m_2} \tilde{e}_i s^i \in \mathcal{S}_1 : \|\vec{s}\| = \sum_i |s^i| \leq 1\}$$

we have a family of E -paths $s \mapsto b_1(s, \vec{s} + \vec{k}_0)$. This family is uniformly regular w.r.t. \vec{s} at $s = 0$ and, moreover, $b_1(0, \vec{s} + \vec{k}_0) = \vec{s} + \vec{k}_0$.

The families $a_{\vec{r}}(s) := \mathbf{d}^{\mathfrak{w}(\vec{r}), p(\vec{r})}(s)$ and $b_{\vec{s}}(s) := b_1(s, \vec{s} + \vec{k}_0)$ satisfy the assumptions of Lemma 8.2, and hence there exist vectors \vec{r}_0, \vec{s}_0 and a number $0 < \varepsilon \leq \eta$ such that

$$[\mathbf{d}^{\mathfrak{w}(\vec{r}_0), p(\vec{r}_0)}(s)]_{s \in [0, \varepsilon]} = [b_1(s, \vec{s}_0 + \vec{k}_0)]_{s \in [0, \varepsilon]}.$$

Projecting equality (9.4) (for $\mathfrak{w} = \mathfrak{w}(\vec{r}_0)$ and $p = p(\vec{r}_0)$) to E and using the above equality, we get

$$\begin{aligned} [b_0(s, p(\vec{r}_0))]_{s \in [0, \varepsilon]} [f(x(t, \varepsilon, p(\vec{r}_0)), u_\varepsilon^{\mathfrak{w}(\vec{r}_0)}(t))]_{t \in [t_0, t_1 + \varepsilon \delta t(\vec{r}_0)]} = \\ [f(x(t), u(t))]_{t \in [t_0, t_1]} [b_1(s, \vec{s}_0 + \vec{k}_0)]_{s \in [0, \varepsilon]}. \end{aligned}$$

Since $b_0(s, p(\vec{r}_0))$ lies in $\text{Im } \Phi_0$ and $b_1(s, \vec{s}_0 + \vec{k}_0)$ in $\text{Im } \Phi_1$, the trajectories $f(x(t), u(t))$, with $t \in [t_0, t_1]$, and $f(x(t, \varepsilon, p(\vec{r}_0)), u_\varepsilon^{\mathfrak{w}(\vec{r}_0)}(t))$, with $t \in [t_0, t_1 + \varepsilon \delta t(\vec{r}_0)]$, are E -homotopic relative to (Φ_0, Φ_1) . On the other hand, from (9.5) we deduce (in the same way as in the proof of Theorem 9.1) that the cost on the first of these trajectories is smaller. This contradicts the optimality of $(\mathbf{x}(t), u(t))$. \square

Now, as a simple consequence of Lemma 9.4, we obtain the following result.

Lemma 9.7. *Let $t \in [t_0, \tau]$ be any regular point of the control u . Then*

$$\mathcal{K}_t^u \subset \text{cl}(\mathcal{K}_\tau^u).$$

Proof. By Lemma 9.4, $\mathcal{K}_t^u \subset \text{cl}(\mathcal{K}_\tau^u)$, and hence

$$\begin{aligned} \mathcal{K}_t^u &= \text{conv}\{\mathcal{B}_{t_1 t_0}(\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)}), \mathcal{K}_t^u\} \subset \text{conv}\{\text{cl}(\mathcal{B}_{t_1 t_0}(\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)})), \text{cl}(\mathcal{K}_\tau^u)\} \\ &= \text{cl}(\text{conv}\{\mathcal{B}_{t_1 t_0}(\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)}), \mathcal{K}_\tau^u\}) = \text{cl}(\mathcal{K}_\tau^u). \end{aligned}$$

□

The above result allows us to define a convex cone

$$\mathcal{K}_{t_1}^u := \bigcup_{\tau < t_1, \tau \text{ regular}} \text{cl}(\mathcal{K}_\tau^u)$$

in $\mathcal{A}_{\underline{x}(t_1)}$. Similarly as in Section 9.1 we have the following result.

Lemma 9.8. *The the convex cones $\mathcal{K}_{t_1}^u$ and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ are separable.*

Proof. If $\mathcal{K}_{t_1}^u$ and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ were not separable, then, by repeating the argument from the proof of Lemma 9.5, the cones $\text{cl}(\mathcal{K}_t^u)$ and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ would not be separable for some regular $t < t_1$. Consequently, also \mathcal{K}_t^u and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$ would not be separable (Lemma D.3), which contradicts Theorem 9.6. □

Now we can finish the proof of Theorem 5.4. From the previous lemma we can deduce that there exists a covector $\xi(t_1) \in \mathcal{A}_{\underline{x}(t_1)}^*$ separating $\mathcal{K}_{t_1}^u$ and $\mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$; i.e.,

$$\langle d, \xi(t_1) \rangle \leq 0 \leq \langle s, \xi(t_1) \rangle \quad \text{for any } d \in \mathcal{K}_{t_1}^u \text{ and } s \in \mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}.$$

We define $\xi(t) := \mathcal{B}_{t t_1}^* \xi(t_1)$. Now, since $\mathcal{K}_{t_1}^u \subset \mathcal{K}_t^u$ and $\Lambda_{\underline{x}(t_1)} = \theta_{\underline{x}(t_1)} \oplus \Lambda_{\underline{x}(t_1)} \subset \mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$, the covector $\xi(t_1)$ separates also \mathcal{K}_t^u and $\Lambda_{\underline{x}(t_1)}$. Consequently, $\xi(t)$ satisfies the assertion of Theorem 5.3. On the other hand, the separating covector $\xi(t_1)$ must vanish on the linear subspaces $\mathcal{B}_{t_1 t}^u(\mathcal{S}_0 \oplus \theta_{\underline{x}(t_0)}) \subset \mathcal{K}_{t_1}^u$ and $\mathcal{S}_1 \oplus \theta_{\underline{x}(t_1)} \subset \mathcal{S}_1 \oplus \Lambda_{\underline{x}(t_1)}$, which gives additional conditions from Theorem 5.4. □

Differnetial geometry

A.1 Lie groupoids

In this section we give the definition of a Lie groupoid, study some fundamental examples and recall the construction of a Lie algebroid of a Lie groupoid. Later we address some questions about integrability of Lie algebroids. The discussion is based mostly on [Mackenzie, 2005].

Lie groupoids

Definition A.1. A *Lie groupoid* consists of two manifolds: \mathcal{G} (a *groupoid*) and M (a *base*), together with two surjective submersions $\alpha, \beta : \mathcal{G} \rightarrow M$ called the *source* and the *target maps*, a smooth map $\iota : x \mapsto \iota_x, M \rightarrow \mathcal{G}$, called the *inclusion map*, and a smooth partial multiplication $(h, g) \mapsto hg, \mathcal{G} * \mathcal{G} = \{(h, g) \in \mathcal{G} \times \mathcal{G} : \alpha(h) = \beta(g)\} \rightarrow \mathcal{G}$, subject to the following conditions:

- $\alpha(hg) = \alpha(g)$ and $\beta(hg) = \beta(h)$ for all $(h, g) \in \mathcal{G} * \mathcal{G}$;
- the partial multiplication is associative; i.e., $j(hg) = (jh)g$ for all j, h, g such that $(j, h), (h, g) \in \mathcal{G} * \mathcal{G}$;
- $\alpha(\iota_x) = \beta(\iota_x) = x$ for all $x \in M$;
- ι is a two-sided identity; i.e., $g\iota_{\alpha(g)} = \iota_{\beta(g)}g = g$ for all $g \in \mathcal{G}$;
- each $g \in \mathcal{G}$ has a two-sided inverse $g^{-1} \in \mathcal{G}$ such that $\alpha(g^{-1}) = \beta(g)$, $\beta(g^{-1}) = \alpha(g)$, $g^{-1}g = \iota_{\alpha(g)}$, and $gg^{-1} = \iota_{\beta(g)}$.

A *morphism of Lie groupoids* \mathcal{G} and $\tilde{\mathcal{G}}$ is a pair of smooth maps $F : \mathcal{G} \rightarrow \tilde{\mathcal{G}}$ and $f : M \rightarrow \tilde{M}$ which preserves the source and target maps; i.e., $f \circ \alpha = \tilde{\alpha} \circ F$ and $f \circ \beta = \tilde{\beta} \circ F$, and preserves the multiplication; that is, $F(h)F(g) = F(hg)$ for all $(h, g) \in \mathcal{G} * \mathcal{G}$.

From the point of view of category theory, a groupoid may be regarded as a small category with the set of objects equal to M and the set of arrows equal to \mathcal{G} , and such that all the arrows are invertible. The word "Lie" refers to the fact that the groupoid structure is compatible with the smooth structures on M and \mathcal{G} (similarly as a Lie group is a group with

a smooth structure compatible with the multiplication). It is an easy exercise to show that the assumption that α and β are surjective submersions imply that $\mathcal{G} * \mathcal{G} = (\alpha \times \beta)^{-1}(\Delta_M)$ has a smooth structure of a submanifold of $\mathcal{G} \times \mathcal{G}$. Moreover, it follows from the smoothness of the partial multiplication and the properties of α and β that the inverse mapping $g \mapsto g^{-1}$ is a diffeomorphism. Details can be found in [Mackenzie, 2005].

Lie groupoids appear naturally in many situations. A basic example is a Lie group G , with $M = \{\text{pt}\}$ being a single point, trivial α and β , and group multiplication. Another standard example is a *pair groupoid* $\mathcal{G} = M \times M$ of a manifold M , with the source $\alpha(x, y) = y$, the target $\beta(x, y) = x$, and the multiplication $(x, y)(y, z) = (x, z)$.

For a (right) principal G -bundle $G \longrightarrow P \xrightarrow{\pi} M$ we can construct an important example of a *gauge groupoid* $\mathcal{G}_P = P \times P / G$ over M . In \mathcal{G}_P we identify pairs (p, q) and (pg, qg) for all $p, q \in P$ and $g \in G$. The source and target maps are simply $\alpha[(p, q)] = \pi(q)$ and $\beta[(p, q)] = \pi(p)$, and the multiplication reads as $[(p, q)][(q, r)] = [(p, r)]$. For the two extreme cases: $(P = M, G = \{1\})$ and $(P = G, M = \{\text{pt}\})$, \mathcal{G}_P is the pair groupoid $M \times M$ and the Lie group G , respectively. Note that \mathcal{G}_P can be regarded as a pair groupoid $P \times P$ divided by the action of G (all the groupoid data for $P \times P$ is G -equivariant).

For a Lie groupoid \mathcal{G} and $x \in M$ we may define an α -fibre of \mathcal{G} over x

$$\mathcal{G}_x := \{g \in \mathcal{G} : \alpha(g) = x\} \subset \mathcal{G}.$$

Note that \mathcal{G}_x is a closed embedded submanifold of \mathcal{G} . The groupoid \mathcal{G} is called α -*simply connected* if each of its α -fibres is simply connected.

Take now an element $g \in \mathcal{G}$. A *right translation* corresponding to g is $R_g : \mathcal{G}_{\beta(g)} \longrightarrow \mathcal{G}_{\alpha(g)}$ defined simply as $h \mapsto hg$.

A Lie algebroid of a Lie groupoid

In this part we describe the construction of a Lie algebroid of a Lie groupoid basing mostly on [Mackenzie, 2005, Silva & Weinstein, 1999]. The procedure follows closely the standard construction of a Lie group–Lie algebra reduction.

Consider a Lie groupoid \mathcal{G} over M and the right action of \mathcal{G} on itself. Since the right translation $R_g : \mathcal{G}_{\beta(g)} \rightarrow \mathcal{G}_{\alpha(g)}$ is a diffeomorphism of α -fibres (not the whole \mathcal{G}) there is a sense speaking of right-invariant vector fields on \mathcal{G} only for fields tangent to α -fibres.

Denote by $T^\alpha \mathcal{G} := \ker T\alpha \subset T\mathcal{G}$ the distribution tangent to the foliation of \mathcal{G} by α -fibres $\mathcal{G}^\alpha = \{\alpha^{-1}(x) : x \in M\}$. A vector field $X \in \text{Sec}(T^\alpha \mathcal{G})$ is said to be *right-invariant* if $(TR_g)(X(h)) = X(hg)$ for all $(h, g) \in \mathcal{G} * \mathcal{G}$. The set of right-invariant vector fields will be denoted by $\mathfrak{X}_R(\mathcal{G})$. Observe that a right-invariant vector field is uniquely determined by its value along the identity section $\iota(M)$. Indeed, we have $X(g) = (TR_g)X(\iota_{\beta(g)})$. Consequently, we can identify the space $\mathfrak{X}_R(\mathcal{G})$ with the space of sections of the bundle $\mathcal{A}(\mathcal{G}) := T^\alpha \mathcal{G}|_{\iota(M)} \longrightarrow M$

$$\mathfrak{X}_R(\mathcal{G}) \approx \text{Sec}(\mathcal{A}(\mathcal{G}), M).$$

Since $\mathcal{A}(\mathcal{G}) \longrightarrow M$ is a pullback of $T^\alpha \mathcal{G} \longrightarrow \mathcal{G}$ via $\iota : M \longrightarrow \mathcal{G}$,

$$\begin{array}{ccc} \mathcal{A}(\mathcal{G}) & \longrightarrow & T^\alpha \mathcal{G} , \\ \downarrow & & \downarrow \\ M & \xrightarrow{\iota} & \mathcal{G} \end{array}$$

it has a smooth vector bundle structure induced from $T^\alpha \mathcal{G}$. What is more, since the Lie bracket of right-invariant vector fields on \mathcal{G} is again right-invariant, the Lie bracket on $\mathfrak{X}_R(\mathcal{G})$ induces a natural skew-symmetric bilinear bracket $[\cdot, \cdot]$ on sections of $\mathcal{A}(\mathcal{G})$. This bracket satisfies the Liebniz rule (2.1) for $\rho = T\beta|_{\mathcal{A}(\mathcal{G})}$. The bundle $\mathcal{A}(\mathcal{G}) \rightarrow M$, together with $[\cdot, \cdot]$ and ρ , is a Lie algebroid called a *Lie algebroid of a Lie groupoid* \mathcal{G} . Lie algebroids which come from some Lie groupoid by the construction described above are called *integrable*.

Note that the maps $TR_{g^{-1}} : T_g^\alpha \mathcal{G} \rightarrow T_{\iota_\beta(g)}^\alpha \mathcal{G} = \mathcal{A}(\mathcal{G})_{\beta(g)}$ defined point-wise for all $g \in \mathcal{G}$ give rise to a vector bundle map (*reduction map*)

$$\begin{array}{ccc} T^\alpha \mathcal{G} & \xrightarrow{\mathcal{R}} & \mathcal{A}(\mathcal{G}) , \\ T\beta \downarrow & & \downarrow \rho \\ TM & \xrightarrow{\text{id}_{TM}} & TM \end{array} \quad (\text{A.1})$$

which is a fibre-wise isomorphism.

The construction of $\mathcal{A}(\mathcal{G})$ can be repeated also for the left action of \mathcal{G} and left-invariant vector fields. It is a matter of convention which construction we use.

Natural examples of Lie algebroids are, in fact, associated with Lie groupoids. A Lie algebroid associated with a Lie group G is its Lie algebra $\mathfrak{g} \approx T_e G$. For the pair groupoid $M \times M$, α -fibres are simply $\alpha^{-1}(x) = M \times \{x\}$, and hence $T_x^\alpha \mathcal{G} \approx T_x M$. A Lie algebroid associated with this groupoid is the *tangent algebroid* on $TM \longrightarrow M$. A Lie algebroid associated with the gauge groupoid \mathcal{G}_P is called an *Atiyah algebroid*. It will be described in detail in Section A.2.

We have shown above that to every Lie groupoid \mathcal{G} corresponds a Lie algebroid $\mathcal{A}(\mathcal{G})$. In fact, also every Lie groupoid morphism $F : \mathcal{G} \longrightarrow \tilde{\mathcal{G}}$ over $f : M \longrightarrow \tilde{M}$ induces a natural morphism of Lie algebroids $\mathcal{A}(F) : \mathcal{A}(\mathcal{G}) \longrightarrow \mathcal{A}(\tilde{\mathcal{G}})$ over $f : M \longrightarrow \tilde{M}$ which can be described as follows. Since F preserves the source map, the vector bundle morphism $TF : T\mathcal{G} \longrightarrow T\tilde{\mathcal{G}}$ restricts to

$$\begin{array}{ccc} T^\alpha \mathcal{G} & \xrightarrow{T^\alpha F} & T^\alpha \tilde{\mathcal{G}} . \\ \downarrow & & \downarrow \\ \mathcal{G} & \xrightarrow{F} & \tilde{\mathcal{G}} \end{array}$$

Now, since $\tilde{\iota} \circ F = f \circ \iota$, the derivative $T^\alpha F$ induces a map of pullbacks $\mathcal{A}(F) : \mathcal{A}(\mathcal{G}) = \iota^* T^\alpha \mathcal{G} \longrightarrow \mathcal{A}(\tilde{\mathcal{G}}) = \tilde{\iota}^* T^\alpha \tilde{\mathcal{G}}$ over $f : M \longrightarrow \tilde{M}$. One can check that this is a morphism of Lie algebroids. Lie algebroid morphisms of the form $\mathcal{A}(F)$ are called *integrable*. In fact the association of a Lie algebroid $\mathcal{A}(\mathcal{G})$ to a Lie groupoid \mathcal{G} , and a Lie algebroid morphism $\mathcal{A}(F)$ to a morphism of Lie groupoids F is a functor from the category of Lie groupoids to the category of Lie algebroids. In literature it is known as a *Lie functor*.

Lie theory

Since a Lie groupoid–Lie algebroid reduction can be considered as a generalisation of the Lie group–Lie algebra reduction, it is natural to ask a question about possible extension of Lie integrability theorems to this new context. This topic is extensively treated in [Mackenzie, 2005] and solved by [Crainic & Fernandes, 2003]. In this work we are interested in two problems:

- Integration of Lie algebroids—does every Lie algebroid is a Lie algebroid of some Lie groupoid?
- Integration of Lie algebroids morphism—can we lift a morphism of two integrable Lie algebroids to a morphism of the corresponding groupoids?

The answer to the first problem is in general negative. A complete solution was given by [Crainic & Fernandes, 2003] (see also [Almeida, 1980, Almeida & Kumpera, 1981, Almeida & Molino, 1987, Cattaneo & Felder, 2004]). The idea goes back to covering theory. Recall that the universal cover \tilde{X} of a topological space X can be constructed as a space of homotopy classes of paths emerging from a fixed point $x_0 \in X$. If $X = G$ is a Lie group, each sufficiently regular path in G can be reduced to a path in \mathfrak{g} —its Lie algebra. Now it turns out that the homotopies of paths in G can be reduced to homotopies in \mathfrak{g} , which are expressed entirely in terms of the Lie algebra structure, without referring to the structure of the underlying Lie group. Consequently, the universal cover \tilde{G} of G can be defined as the space of paths in \mathfrak{g} divided by the equivalence relation coming from homotopy. The group structure on \tilde{G} is given by the composition of paths [Duistermaat & Kolk, 2000]. The same construction can be repeated for Lie algebroids, yet we have to use admissible paths and algebroid homotopies (see Chapter 3). The quotient space, with the multiplication defined by the composition of admissible paths, has a structure of an α -simply connected topological groupoid. Unlike to the case of a Lie algebra, there may be some obstructions to introduce a smooth structure on this groupoid. These are described in [Crainic & Fernandes, 2003]. Theorem 3.6 is closely related with the ideas sketched above. Note, however, that we work in a measurable category, whereas Crainic and Fernandes use the smooth data.

The second integrability problem has a positive solution under mild topological assumptions.

Theorem A.2 ([Mackenzie & Xu, 2000]). *Consider Lie groupoids \mathcal{H} over S and \mathcal{G} over M , and suppose that $\Phi : \mathcal{A}(\mathcal{H}) \rightarrow \mathcal{A}(\mathcal{G})$ over $f : S \rightarrow M$ is an algebroid morphism. If \mathcal{H} is α -simply connected, there is a unique morphism of Lie groupoids $F : \mathcal{H} \rightarrow \mathcal{G}$ over $f : S \rightarrow M$ such that $\mathcal{A}(F) = \Phi$.*

We shall now study groupoid morphisms $\mathcal{H} \rightarrow \mathcal{G}$ in a special situation when $\mathcal{H} = S \times S$ is a pair groupoid. From Theorem A.2 we can easily derive the following result.

Lemma A.3. *Let $\Phi : TS \rightarrow \mathcal{A}(\mathcal{G})$ over $f : S \rightarrow M$ be a morphism of Lie algebroids. Then, if S is connected, there is at most one morphism of Lie groupoids $F : S \times S \rightarrow \mathcal{G}$ integrating Φ . If S is simply connected then such a morphism F exists.*

Proof. If S is simply connected, the existence of F follows immediately from Theorem A.2, as the groupoid $S \times S$ is α -simply connected (α -fibres are of the form $S \times \{x_0\}$).

Assume now that S is an arbitrary connected manifold and let $F : S \times S \rightarrow \mathcal{G}$ be a morphism integrating Φ . Consider the universal cover $\pi : (\tilde{S}, \tilde{x}_0) \rightarrow (S, x_0)$. Clearly, $\tilde{S} \times \tilde{S} \xrightarrow{\pi \times \pi} S \times S$ is a groupoid morphism, hence also the composition $\tilde{S} \times \tilde{S} \xrightarrow{\pi \times \pi} S \times S \xrightarrow{F} \mathcal{G}$ is a groupoid morphism. Moreover, it integrates the Lie algebroid morphism $T\tilde{S} \xrightarrow{T\pi} TS \xrightarrow{\Phi} \mathcal{A}(\mathcal{G})$. We will prove that uniqueness of $F \circ (\pi \times \pi)$ (following from Theorem A.2) implies the uniqueness of F . To see this, observe that if $F \circ (\pi \times \pi) = F' \circ (\pi \times \pi)$ then the set of points in $S \times S$ on which F and F' coincide would be: closed, since F and F' are continuous; nonempty, since $F(x_0, x_0) = F'(x_0, x_0) = \iota_{f(x_0)}$; and open, since $\pi \times \pi$ is a covering. We deduce that it is the whole $S \times S$. \square

Observe now that, if $F : S \times S \rightarrow \mathcal{G}$, with S connected, is a groupoid morphism integrating $\Phi : TS \rightarrow \mathcal{A}(\mathcal{G})$ over $f : S \rightarrow M$, then, for each $x_0 \in S$, $\tilde{\Phi} : x \mapsto F(x, x_0)$ is a map from S to the α -fibre $\mathcal{G}_{f(x_0)}$ of \mathcal{G} such that the following diagram of vector bundle morphisms commutes

$$\begin{array}{ccc} TS & \xrightarrow{T\tilde{\Phi}} & T(\mathcal{G}_{f(x_0)}) \hookrightarrow T^\alpha \mathcal{G} \\ & \searrow \Phi & \downarrow \mathcal{R} \\ & & \mathcal{A}(\mathcal{G}) \end{array} \quad (\text{A.2})$$

Note that, if $\tilde{\Phi} : S \rightarrow \mathcal{G}_{f(x_0)}$ is a smooth map such that (A.2) is satisfied, then $F_{\tilde{\Phi}}(x, y) := \tilde{\Phi}(x)^{-1}\tilde{\Phi}(y)$ is a groupoid morphism $F_{\tilde{\Phi}} : S \times S \rightarrow \mathcal{G}$ integrating Φ .

By Lemma A.3, $F_{\tilde{\Phi}}$ is unique. Consequently, if $\tilde{\Phi}, \tilde{\Phi}' : S \rightarrow \mathcal{G}_{f(x_0)}$ are two maps satisfying (A.2), they are related by $\tilde{\Phi}(x)^{-1}\tilde{\Phi}(y) = \tilde{\Phi}'(x)^{-1}\tilde{\Phi}'(y)$, and hence $\tilde{\Phi}'(y) = \tilde{\Phi}'(x)\tilde{\Phi}(x)^{-1}\tilde{\Phi}(y)$.

Corollary A.4. *Let S be a connected manifold. If a Lie algebroid morphism $\Phi : TS \rightarrow \mathcal{A}(\mathcal{G})$ over $f : S \rightarrow M$ is integrable then, for each $x_0, y_0 \in S$ and $g \in \alpha^-(f(x_0)) \cap \beta^{-1}(f(y_0))$, there exists a unique smooth map $\tilde{\Phi} : S \rightarrow \mathcal{G}_{f(x_0)} = \alpha^{-1}(f(x_0))$ such that $\mathcal{R} \circ T\tilde{\Phi} = \Phi$ and $\tilde{\Phi}(y_0) = g$.*

Conversely, if $\tilde{\Phi}$ as above exists (for some x_0, y_0 and g), then the Lie algebroid morphism Φ is integrable. In particular, such $\tilde{\Phi}$ exists if S is simply connected.

A.2 The Atiyah algebroid

In this section we describe the Atiyah algebroid—a Lie algebroid canonically associated with a principal bundle. In particular we study its Lie bracket and give a description of the associated linear Poisson structure. Our discussion is based mostly on [Mackenzie, 2005].

Invariant vector fields

Definition A.5. A *principal bundle* $G \longrightarrow P \xrightarrow{\pi} M$ is a locally trivial fibre bundle $\pi : P \longrightarrow M$, equipped with a (right) free action of the Lie group G on P ; $(p, g) \mapsto pg$; $P \times G \longrightarrow P$ such that its orbits coincide with the fibres of π .

Observe that if $G \longrightarrow P \xrightarrow{\pi} M$ is a principal bundle, the action of G on P induces the action on the tangent bundle TP . Denote this action by R_g . The quotient space TP/G has a natural structure of a vector bundle over M . We define the addition simply by

$$[X_p] + [Y_{pg}] = [R_g X_p + Y_{pg}]$$

and the base projection by

$$[X_p] \mapsto \pi(p).$$

It is straightforward to verify that the above constructions are well-defined.

Observe that section of the quotient bundle $E := TP/G \rightarrow M$ can be canonically identified with G -invariant vector fields on P . Note that, since G -invariant vector fields on P are closed under the Lie bracket $[\cdot, \cdot]_{TP}$, we have an induced bracket $[\cdot, \cdot]_E$ on the space of sections of E . Clearly, $[\cdot, \cdot]_E$ inherits the skew-symmetry and the Jacobi identity from $[\cdot, \cdot]_{TP}$, and hence $(\text{Sec}(E), [\cdot, \cdot]_E)$ is a Lie algebra. Moreover, since for any G -invariant vector fields $X, Y \in \mathfrak{X}^G(P)$, and every base function $f \in C^\infty(M)$ we have

$$[X, \pi^* f Y]_{TP} = \pi^* f [X, Y]_{TP} + (T\pi)(X)(f)Y,$$

and the derivative $(T\pi)(X)(f)$ depends only on the class of X , the bracket $[\cdot, \cdot]_E$ satisfies the Leibniz rule

$$[[X], f[Y]]_E = f [[X], [Y]]_E + \rho([X])(f)[Y],$$

where the anchor map $\rho : E \rightarrow TM$ is defined by $\rho([X]) = T\pi(X)$. Clearly, ρ satisfies also the compatibility condition (2.2).

To sum up, the bundle $E = TP/G \longrightarrow M$, together with the bracket $(E, [\cdot, \cdot]_E, \rho)$, and the anchor map $\rho : E \longrightarrow TM$ is a Lie algebroid.

Definition A.6. The Lie algebroid structure on $\pi : TP/G \longrightarrow M$ described above is called an *Atiyah algebroid* of the principal bundle $G \longrightarrow P \xrightarrow{\pi} M$.

As has been already mentioned the Atiyah algebroid can be also described as a Lie algebroid associated with the gauge groupoid \mathcal{G}_P . Now, we shall investigate this structure in detail.

The Atiyah sequence

Observe that, since G acts on the fibres of π , the action R_g restricts to the space $VP \subset TP$ of vertical vectors (i.e., vectors tangent to the fibres of π). Since VP is spanned by the fundamental vector fields of the G -action on P , we have a canonical isomorphism

$VP \simeq P \times \mathfrak{g}$, where \mathfrak{g} is the Lie algebra of G . The action R_g in this identification reads $R_g(p, a) = (pg, Ad_{g^{-1}}a)$, so $K := VP/G \simeq P \times_G \mathfrak{g}$, where the right action of G on \mathfrak{g} is $g \mapsto Ad_{g^{-1}}$. What is more, the bracket of two G -invariant vertical vector fields on P corresponds, in this identification, to the canonical (right) Lie bracket $[\cdot, \cdot]_{\mathfrak{g}}$ on \mathfrak{g} . That is, if $x \mapsto (x, a(x))$ and $x \mapsto (x, b(x))$ are two G -invariant sections of $P \times \mathfrak{g} \simeq VP$, then

$$[(x, a(x)), (x, b(x))]_{TP} = (x, [a(x), b(x)]_{\mathfrak{g}}) \in P \times \mathfrak{g}.$$

This shows that the bundle K is a Lie algebroid with the trivial anchor and the Lie algebra structure in fibres isomorphic to \mathfrak{g} . Alternatively, we may argue that the sections of K can be identified with G -invariant vertical vector fields on P , which are closed under the Lie bracket. The Lie algebroid structure on K is thus the restriction of the Lie algebroid structure on TP/G to vertical vector fields. Hence, we get the following (exact) sequence of Lie algebroid morphisms called the *Atiyah sequence*:

$$0 \rightarrow K := P \times_G \mathfrak{g} \rightarrow E \xrightarrow{\rho} TM \rightarrow 0.$$

Local description

Introduce now a local trivialisation $\phi_s : V \times G \rightarrow P|_V$ obtained from a local section $s : V \rightarrow P$ by the formula $\phi_s(x, g) = s(x)g$. Clearly, we may identify $E|_V$ with $TP|_{s(V)}$, thus $E|_V \simeq TM|_V \times \mathfrak{g}$. For two local sections $x \mapsto (X(x), a(x))$ and $x \mapsto (Y(x), b(x))$ in this trivialisation the Lie bracket reads as

$$[(X, a), (Y, b)]_E = ([X, Y]_{TM}, [a, b]_{\mathfrak{g}} + X(b) - Y(a)) \quad (\text{A.3})$$

and the anchor is $\rho((X, a)) = X$. In fact, this is just the product of the Lie algebroids $TM|_V \rightarrow V$ and \mathfrak{g} (cf. the last paragraph of Chapter 2). We have a similar description globally if the principal bundle P is trivial. Note, however, that, if we will work with principal bundles over a neighbourhood of a path in M , we can always assume that our principal bundle is trivial.

In some applications one has to work with a *principal connection* on P . It corresponds to a G -invariant horizontal distribution in TP and is represented by a splitting $E = TM \oplus_M K$ given by a bundle embedding $\nabla : TM \rightarrow E$ such that $\rho \circ \nabla = \text{id}_{TM}$. If the bundle K is trivial, then we get another trivialisation $E \simeq_{\nabla} TM \times \mathfrak{g}$, associated with the connection ∇ , in which the Lie bracket reads as

$$[(X, a), (Y, b)]_E = ([X, Y]_{TM}, F_{\nabla}(X, Y) + [a, b]_{\mathfrak{g}} + X(b) - Y(a)); \quad (\text{A.4})$$

where F_{∇} is the *curvature* of the connection ∇ , i.e.,

$$F_{\nabla}(X, Y) = [\nabla(X), \nabla(Y)]_E - \nabla([X, Y]_{TM}).$$

The anchor map is still simply $\rho((X, a)) = X$.

Observe that formula (A.3) is a special case of (A.4). Indeed, we may regard a local section $s : V \rightarrow P$ as a G -invariant horizontal local distribution on P obtained by spanning $Ts \subset TP$ by the G -action. Clearly the curvature of this distribution vanishes, and hence (A.4) and (A.3) coincide in this special case.

The linear Poisson structure

Now we shall describe the linear Poisson structure Π_E on the dual bundle $E^* = T^*P/G \longrightarrow M$ canonically associated with the Lie algebroid structure on TP/G .

Suppose for simplicity that the vertical subbundle $K = VP/G$ is trivial (e.g. P is trivial), $K = M \times \mathfrak{g}$, and consider a splitting $E \simeq_{\nabla} TM \oplus_M K$ induced by a principal connection $\nabla : TM \rightarrow E$, so that we get an identification $E \simeq_{\nabla} TM \times \mathfrak{g}$. Let $E^* \simeq_{\nabla} T^*M \times \mathfrak{g}^*$ be the corresponding identification of the dual bundle.

Theorem A.7. *The Poisson tensor Π_E associated with the Lie algebroid structure on E in the identification $E^* \simeq_{\nabla} T^*M \times \mathfrak{g}^*$ reads as*

$$\Pi_E(p_x, \zeta) = \left(\Pi_{T^*M}(p_x) + V_{\langle \zeta, F_{\nabla}(x)(\cdot, \cdot) \rangle} \right) \times \Pi_{\mathfrak{g}^*}(\zeta),$$

where $p_x \in T_x^*M$, $\xi \in \mathfrak{g}^*$, Π_{T^*M} and $\Pi_{\mathfrak{g}^*}$ are the standard Poisson tensors, and $V_{\langle \zeta, F_{\nabla}(x)(\cdot, \cdot) \rangle}$ is the two-form $\langle \zeta, F_{\nabla}(x)(\cdot, \cdot) \rangle$ associated with the curvature $F_{\nabla}(x) : \wedge^2 T_x M \rightarrow \mathfrak{g}$ understood as a vertical tensor on T^*M .

Consequently, the Hamiltonian vector field defined by means of Π_E and a Hamiltonian $h : T^*M \times \mathfrak{g}^* \longrightarrow \mathbb{R}$ reads as

$$\mathcal{X}_h^E(p_x, \zeta) = \left(\mathcal{X}_{h(\cdot, \zeta)}^{T^*M}(p_x) + V_{\langle \zeta, F_{\nabla}(x)(\frac{\partial h}{\partial p}(p_x, \zeta), \cdot) \rangle}, \mathcal{X}_{h(p_x, \cdot)}^{\mathfrak{g}^*}(\zeta) \right),$$

where with $\mathcal{X}_h^{T^*M}$ and $\mathcal{X}_h^{\mathfrak{g}^*}$ we denoted the Hamiltonian vector fields associated with Poisson structures Π_{T^*M} and $\Pi_{\mathfrak{g}^*}$.

In local coordinates, $p \sim (x^a, p_b)$ and $\zeta \sim (\zeta_\alpha)$,

$$\begin{aligned} \mathcal{X}_h^E(x, p, \zeta) = & \frac{\partial h}{\partial p_a}(x, p, \zeta) \partial_{x^a} + \left(\zeta_\alpha F_{ab}^\alpha(x) \frac{\partial h}{\partial p_a}(x, p, \zeta) - \frac{\partial h}{\partial x^b}(x, p, \zeta) \right) \partial_{p_b} \\ & + \zeta_\gamma C_{\alpha\beta}^\gamma \frac{\partial h}{\partial \zeta_\alpha} \partial_{\zeta_\beta}, \end{aligned}$$

where $F_{ab}^\alpha(x)$ are the coefficients of the curvature F_{∇} and $C_{\beta\gamma}^\alpha$ are the structure constants of \mathfrak{g} .

Proof. The proof is straightforward by an explicit coordinate calculation using the local description of the bracket and the anchor map (A.4) and formula (2.3) relating the bracket and anchor with the coefficient of the linear Poisson tensor. \square

Analysis

B.1 Mesuaralbe maps, regular points

In this section we briefly recall some basic properties of measurable and absolutely continuous maps. Later we introduce a notion of *uniform regularity* and study its basic properties. It is a quite important technical tool in our considerations.

Basic facts

When speaking about measure we will always have in mind *Lebesgue measure* in \mathbb{R}^n or subsets of \mathbb{R}^n . This measure will be denoted by $\mu_L(\cdot)$.

Recall that a map $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$, defined on a subset $V \subset \mathbb{R}^n$, is *measurable* if the inverse image of every open set is Lebesgue measurable in V . The measurable map f will be called *bounded* if the closure of its image is a compact set. Observe that every bounded (or locally bounded) measurable function is locally integrable.

Measurable maps can be characterised as follows.

Theorem B.1 (Luzin). *The map $f : \mathbb{R}^n \supset V \rightarrow \mathbb{R}$ defined on a measurable set V is measurable iff, for every $\varepsilon > 0$, there exists a closed subset $F \subset V$ such that the restriction $f|_F$ is continuous and $\mu_L(V \setminus F) < \varepsilon$.*

For the proof see [Łojasiewicz & Ferreira, 1988].

In our considerations much attention will be paid to regular points of measurable maps.

Definition B.2. Let $f : [a, b] \rightarrow \mathbb{R}^m$ be a measurable map. A point $x \in [a, b]$ is called a *regular point* (also: *Lebesgue* or *density point*) of f , iff

$$\lim_{t \rightarrow 0} \frac{1}{|t|} \int_0^t |f(x+s) - f(x)| ds = 0.$$

For bounded measurable (or more generally integrable) maps we have the following result.

Theorem B.3 (Lebesgue). *For an integrable map $f : [a, b] \rightarrow \mathbb{R}^m$ almost every point in $[a, b]$ is a regular point of f .*

For the proof see [Łojasiewicz & Ferreira, 1988]

A map $x : [a, b] \rightarrow \mathbb{R}^k$ is called *absolutely continuous* (AC) if it can be written in the form

$$x(t) = x(a) + \int_a^t v(\tau) d\tau,$$

where $v(\cdot)$ is an integrable map. As we see, an AC map $x(t)$ is differentiable at all the regular points t of v (hence, by Lebesgue Theorem, differentiable a.e.). Its derivative at such a point is simply $v(t)$. In this work we concentrate our attention mostly on *absolutely continuous maps with bounded derivative* (ACB maps) i.e. maps for which $v(\cdot)$ is bounded measurable.

In our considerations we will use the following lemma.

Lemma B.4. *Let $a : [0, 1] \rightarrow \mathbb{R}$ be a bounded measurable map, let $h : [0, 1] \rightarrow [0, 1]$ be a continuous function, and let $g : [0, 1] \rightarrow [0, 1]$ be a C^1 -map with a non-vanishing derivative. Then the map*

$$s \mapsto G(s) := \int_0^1 |a(g(t)h(s)) - a(g(t)h(s_0))| dt$$

is regular (in fact continuous) at every s_0 such that $h(s_0) \neq 0$.

Proof. Let c be a number such that $0 < c \leq |g'(x)|$ for every $x \in [0, 1]$. Now if $A \subset [0, 1]$ is a measurable subset then $\mu_L(g^{-1}(A)) \leq \frac{1}{c}\mu_L(A)$.

Choose $\varepsilon > 0$. By Luzin Theorem B.1 there exists a closed set $F \subset [0, 1]$ such that $a(\cdot)$ is continuous on F and $\mu_L([0, 1] \setminus F) < \varepsilon$. Now a is uniformly continuous on F , g is bounded and h continuous, hence there exists $\delta > 0$ such that $|a(g(t)h(s)) - a(g(t)h(s_0))| < \varepsilon$ if only $|s - s_0| < \delta$ and t and s are such that $g(t)h(s) \in F$ and $g(t)h(s_0) \in F$. As a consequence for $|s - s_0| < \delta$, we can estimate

$$\begin{aligned} \int_0^1 |a(g(t)h(s)) - a(g(t)h(s_0))| dt &\leq \int_{\{t: g(t)h(s) \notin F\}} 2\|a\| dt + \int_{\{t: g(t)h(s_0) \notin F\}} 2\|a\| dt + \\ &+ \int_{\{t: g(t)h(s) \in F\} \cap \{t: g(t)h(s_0) \in F\}} |a(g(t)h(s)) - a(g(t)h(s_0))| dt \\ &\leq 2\|a\| \cdot \mu_L\left(g^{-1}\left(\frac{1}{h(s)}([0, 1] \setminus F)\right)\right) + 2\|a\| \cdot \mu_L\left(g^{-1}\left(\frac{1}{h(s_0)}([0, 1] \setminus F)\right)\right) + \int_0^1 \varepsilon dt \\ &\leq \varepsilon \left(2\|a\| \frac{1}{c} \left(\frac{1}{h(s)} + \frac{1}{h(s_0)}\right) + 1\right) \end{aligned}$$

Since $h(s_0) \neq 0$, if $|s - s_0|$ is sufficiently small, the values s of $G(s)$ are arbitrarily close to $0 = G(s_0)$, which finishes the proof. \square

Uniform regularity

Regular points play an important role in our considerations, since the behaviour of a measurable map at a regular point is similar to the behaviour of a continuous map. To study behaviour of the families of measurable maps we introduce a notion of *uniform regularity*.

Definition B.5. Let P be a topological space and consider a map $f : [a, b] \times P \rightarrow \mathbb{R}^m$ such that $t \mapsto f(t, p)$ is measurable for every $p \in P$. We call f *uniformly regular with respect to* $p \in P$ at $x \in [a, b]$ iff the following conditions are satisfied:

$$\frac{1}{|t|} \int_0^t |f(x+s, p) - f(x, p)| \, ds \xrightarrow[t \rightarrow 0]{} 0 \quad \text{locally uniformly w.r.t. } p, \quad (\text{B.1})$$

$$\text{the map } p \mapsto f(x, p) \text{ is continuous,} \quad (\text{B.2})$$

and for every compact set $K \subset P$ there exists a number $t_0 > 0$ such that

$$p \mapsto \left([0, t_0] \ni s \mapsto f(x+s, p) \right) \text{ is a continuous map from } K \text{ to } L^1([0, t_0], \mathbb{R}^m). \quad (\text{B.3})$$

Usually in mathematics the word "uniform" means "in the same way for all parameters". In the context of regularity this can be expressed by the condition (B.1) itself. Therefore Definition B.5 is more specific than what one could expect under the name "uniform regularity". The sense of this definition is, however, to abstract several technical properties of measurable maps which are important from the point of view of this work. Since, according to our knowledge, the notion of uniform regularity is not a well established term, we hope that Definition B.5 would not be confusing.

Let us now investigate some simple properties of uniformly regular maps. In what follows we will consider only uniform regularity at point $0 \in \mathbb{R}$ and restrict our attention to parameter spaces P which are metric (we can think of P as of a subset of \mathbb{R}^m).

A basic example of a uniformly regular map is just a continuous map.

Proposition B.6. *Let $F : \mathbb{R} \times P \rightarrow \mathbb{R}$ be a continuous map. Then F is uniformly regular w.r.t. $p \in P$ at $s = 0$.*

Proof. Condition (B.2) is obvious. Fix now a compact set $K_P \subset P$ and restrict s to a fixed interval $[0, t_0]$. Since F is uniformly continuous on $[0, t_0] \times K_P$, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|F(s, p) - F(0, p)| < \varepsilon$ if $|s| < \delta$ and for all $p \in K_P$. Consequently,

$$\int_0^t |F(s, p) - F(0, p)| \, ds \leq |t| \varepsilon$$

for $|t| < \delta$ and all $p \in K_P$. This proves (B.2), i.e., $\frac{1}{|t|} \int_0^t |F(s, p) - F(0, p)| \, ds \xrightarrow[t \rightarrow 0]{} 0$ uniformly w.r.t. $p \in K_P$.

To check (B.3) observe that, by the uniform continuity of F on $[0, t_0] \times K_P$, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|F(s, p) - F(s, p')| \leq \varepsilon$ for every $s \in [0, t_0]$ and all $p, p' \in K_P$ such that $|p - p'| < \delta$. Consequently,

$$\int_0^{t_0} |F(s, p) - F(s, p')| \, ds \leq |t_0| \varepsilon$$

for all $p, p' \in K_P$ such that $|p - p'| < \delta$. This proves (B.3). \square

Another simple example is the following.

Proposition B.7. *Let $f : \mathbb{R} \rightarrow \mathbb{R}^m$ be a measurable map regular at $s = 0$. For $p \in P$ define $\tilde{f}(s, p) := f(s)$. Then the map \tilde{f} is uniformly regular w.r.t. $p \in P$ at $s = 0$.*

Proof. Conditions (B.1)–(B.3) are trivially satisfied. \square

Below we discuss several ways of generating uniformly regular maps from given ones.

Proposition B.8. *Let $f, g : \mathbb{R} \times P \rightarrow \mathbb{R}^m$ be two maps uniformly regular w.r.t. $p \in P$ at $s = 0$. Then the sum $f + g$ is also uniformly regular w.r.t. $p \in P$ at $s = 0$.*

Proof. Property (B.1) is clear since

$$|f(s, p) + g(s, p) - f(0, p) - g(0, p)| \leq |f(s, p) - f(0, p)| + |g(s, p) - g(0, p)|.$$

Property (B.2) is obvious as the sum of continuous maps is continuous.

To prove (B.3) fix a compact set $K \subset P$ and assume that $K \ni p \mapsto ([0, t_0] \ni s \mapsto f(s, p))$ and $K \ni p \mapsto ([0, \tilde{t}_0] \ni s \mapsto g(s, p))$ are continuous. Without loss of generality $t_0 \leq \tilde{t}_0$. Now the restriction $K \ni p \mapsto ([0, t_0] \ni s \mapsto g(s, p))$ is also continuous since

$$\|a(\cdot)\|_{L^1([0, t_0], \mathbb{R}^m)} \leq \|a(\cdot)\|_{L^1([0, \tilde{t}_0], \mathbb{R}^m)}.$$

Consequently, $K \ni p \mapsto f(\cdot, p) + g(\cdot, p) \in L^1([0, t_0], \mathbb{R}^m)$ is continuous as a sum of two continuous maps. \square

Proposition B.9. *Let $f : \mathbb{R} \times P \rightarrow \mathbb{R}^m$ be bounded and uniformly regular w.r.t. $p \in P$ at $s = 0$. Let $h : P \rightarrow \mathbb{R}$ be a continuous map. Then the map $\tilde{f}(s, p) = h(p)f(s, p)$ is uniformly regular w.r.t. p at $s = 0$.*

Proof. Condition (B.2) is obvious. Choose now a compact set $K \subset P$. For $p \in K$ we have

$$\frac{1}{|t|} \int_0^t |h(p)f(s, p) - h(p)f(0, p)| ds \leq \sup_{p \in K} |h(p)| \cdot \frac{1}{|t|} \int_0^t |f(s, p) - f(0, p)| ds \xrightarrow[t \rightarrow 0]{} 0$$

uniformly w.r.t. $p \in K$, and hence (B.1) is satisfied.

Finally, note that for $p, p' \in K$ we have

$$\begin{aligned} & \int_0^{t_0} |h(p)f(s, p) - h(p')f(s, p')| ds \\ & \leq |h(p)| \int_0^{t_0} |f(s, p) - f(s, p')| ds + |h(p) - h(p')| \int_0^{t_0} |f(s, p')| ds \\ & \leq \sup_{p \in K} |h(p)| \int_0^{t_0} |f(s, p) - f(s, p')| ds + |h(p) - h(p')| \|f(\cdot, p')\|_{L^1} \xrightarrow[p \rightarrow p']{} 0 + 0; \end{aligned}$$

that is, (B.3) is satisfied. \square

Lemma B.10. *Let $f : \mathbb{R} \times P \rightarrow \mathbb{R}^m$ be bounded and uniformly regular w.r.t. $p \in P$ at $s = 0$. Consider $\tilde{f}(s, p, c) := f(sc, p)$ where $c \in \mathbb{R}$. Then $\tilde{f} : \mathbb{R} \times P \times \mathbb{R} \rightarrow \mathbb{R}^m$ is uniformly regular w.r.t. $p \in P$ and $c \in \mathbb{R}$ at $s = 0$.*

Proof. Since $\tilde{f}(0, p, c) = f(0, p)$, condition (B.2) is obvious.

Consider now compact sets $K_C \subset \mathbb{R}$ and $K_P \subset P$. For $c \in K_C$ and $p \in K_P$ we have

$$\begin{aligned} \frac{1}{|t|} \int_0^t \left| \tilde{f}(s, p, c) - \tilde{f}(0, p, c) \right| ds &= \frac{1}{|t|} \int_0^t |f(sc, p) - f(0, p)| ds \\ &= \frac{1}{|t|c} \int_0^{tc} \left| f(s', p) - f(0, p) \right| ds' \xrightarrow{t \rightarrow 0} 0. \end{aligned}$$

Since K_C is bounded and $\frac{1}{|t'|} \int_0^{t'} |f(s, p) - f(0, p)| ds \xrightarrow{t' \rightarrow 0} 0$ uniformly w.r.t. $p \in K_P$, the above convergence is uniform w.r.t. $p \in K_P$ and $c \in K_C$.

We are left with the proof of property (B.3). We will check that $(p, c) \mapsto \tilde{f}(\cdot, p, c); K_P \times K_C \rightarrow L^1([0, \tilde{t}_0], \mathbb{R}^m)$ is continuous separately w.r.t. p and w.r.t. c for a suitably chosen \tilde{t}_0 .

Let $t_0 > 0$ be a number from the property (B.3) for $f(s, p)$ and $K = K_P$. To prove the continuity w.r.t. p fix $c \in K_C$. If $c \neq 0$, then

$$\begin{aligned} \int_0^t \left| \tilde{f}(s, p, c) - \tilde{f}(s, p', c) \right| ds &= \int_0^t \left| f(sc, p) - f(sc, p') \right| ds \\ &= \frac{1}{c} \int_0^{tc} \left| f(s', p) - f(s', p') \right| ds' \xrightarrow{p \rightarrow p'} 0 \end{aligned}$$

if only $tc \leq t_0$.

For $c = 0$ we have

$$\int_0^t \left| \tilde{f}(s, p, c) - \tilde{f}(s, p', c) \right| ds = \int_0^t \left| f(0, p) - f(0, p') \right| ds = |t| \left| f(0, p) - f(0, p') \right| \xrightarrow{p \rightarrow p'} 0$$

for every t . In particular, we proved continuity w.r.t. p for $\tilde{t}_0 := \frac{t_0}{\sup_{c \in K_C} |c|}$.

Now fix $p \in P$, fix $c' \in K_C$, choose $\varepsilon > 0$ and consider $c \in K_C$. If $c' = 0$, then

$$\begin{aligned} \int_0^{\tilde{t}_0} \left| \tilde{f}(s, p, c) - \tilde{f}(c, p, c') \right| ds \\ = \int_0^{\tilde{t}_0} |f(sc, p) - f(0, p)| ds = |\tilde{t}_0| \frac{1}{|\tilde{t}_0|c} \int_0^{\tilde{t}_0 c} \left| f(s', p) - f(0, p) \right| ds' \xrightarrow{c \rightarrow 0} 0. \end{aligned}$$

If $c' \neq 0$ consider a closed set $F \subset [0, \tilde{t}_0]$ such that $\mu_L([0, \tilde{t}_0] \setminus F) < \varepsilon |\tilde{t}_0|$ and $\tilde{f}(\cdot, p, c')$ is continuous on F (note that p and c' are fixed). Such a set exists by Luzin Theorem B.1.

Since $\tilde{f}(\cdot, p, c')$ is uniformly continuous on F , there exists a number $\delta > 0$ such that $\left| \tilde{f}(s, p, c') - \tilde{f}(s', p, c') \right| < \varepsilon$ if $|s - s'| \leq \delta$ and $s, s' \in F$. Now $\tilde{f}(s, p, c) = \tilde{f}(\frac{c}{c'}s, p, c')$ and $\left| s - \frac{c}{c'}s \right| \leq \left| \frac{c' - c}{c'} \right| |\tilde{t}_0|$ for $s \in [0, \tilde{t}_0]$, so we have

$$\left| \tilde{f}(s, p, c') - \tilde{f}(s, p, c) \right| < \varepsilon \quad \text{if} \quad |c' - c| \leq \frac{c'}{|\tilde{t}_0|} \delta \quad \text{and} \quad s \in F \cap \frac{c'}{c} F.$$

Note that

$$\begin{aligned} \mu_L \left([0, \tilde{t}_0] \setminus F \cap \frac{c'}{c} F \right) &\leq \mu_L ([0, \tilde{t}_0] \setminus F) + \mu_L \left([0, \tilde{t}_0] \setminus \frac{c'}{c} F \right) \\ &\leq \varepsilon |\tilde{t}_0| + \left(\left| 1 - \frac{c'}{c} \right| + \frac{c'}{c} \varepsilon \right) |\tilde{t}_0| \leq 4\varepsilon |\tilde{t}_0| \end{aligned}$$

if $|c - c'|$ is small enough. Consequently,

$$\begin{aligned} &\int_0^{\tilde{t}_0} \left| \tilde{f}(s, p, c) - \tilde{f}(s, p, c') \right| ds \\ &\leq \int_{F \cap \frac{c'}{c} F} \left| \tilde{f}(s, p, c) - \tilde{f}(s, p, c') \right| ds + \int_{[0, \tilde{t}_0] \setminus F \cap \frac{c'}{c} F} \left| \tilde{f}(s, p, c) - \tilde{f}(s, p, c') \right| ds \\ &\leq \varepsilon |\tilde{t}_0| + \mu_L \left([0, \tilde{t}_0] \setminus F \cap \frac{c'}{c} F \right) \cdot 2 \|f\| \leq \varepsilon (|\tilde{t}_0| + 8 |\tilde{t}_0| \|f\|) \end{aligned}$$

if $|c - c'|$ is small enough. Since ε is an arbitrary positive number, this proves the continuity of $(p, c) \mapsto \tilde{f}(\cdot, p, c)$ w.r.t. c . \square

Lemma B.11. *Let $f : \mathbb{R} \times P \rightarrow \mathbb{R}^m$ be bounded and uniformly regular w.r.t. $p \in P$ at $s = 0$, and let $G : \mathbb{R}^m \times \mathbb{R} \times P \times Q \rightarrow \mathbb{R}^m$ be a continuous map w.r.t all variables. Then the composition $G(f(s, p), s, p, q)$ is uniformly regular w.r.t. $p \in P$ and $q \in Q$ at $s = 0$.*

Proof. We will prove the assertion for G trivially depending on $q \in Q$. This will suffice, since we can denote $G(f(s, p), s, p, q)$ as $G(f(s, p, q), s, p, q) = G(f(s, \tilde{p}), s, \tilde{p})$, where $\tilde{f}(s, p, q) := f(s, p)$ and $\tilde{p} := (p, q) \in P \times Q$. Clearly, $f(s, \tilde{p})$ is uniformly regular w.r.t. $\tilde{p} = (p, q) \in P \times Q$ at $s = 0$ (cf. Proposition B.7) and the investigated composition has a desired simpler form $G(\tilde{f}(s, \tilde{p}), s, \tilde{p})$.

Property (B.2) is obvious. To prove (B.1) estimate

$$\begin{aligned} &|G(f(s, p), s, p) - G(f(0, p), 0, p)| \\ &\leq |G(f(s, p), s, p) - G(f(0, p), s, p)| + |G(f(0, p), s, p) - G(f(0, p), 0, p)|. \end{aligned}$$

Since $\tilde{G}(s, p) := G(f(0, p), s, p)$ is continuous w.r.t. p and s , it satisfies (B.1). Consequently, it is enough to check if

$$\frac{1}{|t|} \int_0^t |G(f(s, p), s, p) - G(f(0, p), s, p)| ds \xrightarrow{t \rightarrow 0} 0$$

locally uniformly w.r.t. p . To prove it consider a compact set $K_P \subset P$ and restrict s to the interval $[0, t_0] \subset \mathbb{R}$. Since f is bounded, its image $\text{Im } f$ is contained in a compact subset $K \subset \mathbb{R}^m$. Fix $\varepsilon > 0$. The map G is uniformly continuous on $K \times [0, t_0] \times K_P$, so there exists a number $\delta > 0$ such that

$$|G(x, s, p) - G(y, s, p)| < \varepsilon$$

if $|x - y| < \delta$ and $x, y \in K$, $s \in [0, t_0]$, and $p \in K_P$. Since $f(s, p)$ is uniformly regular there exists a number $0 < \tilde{t}_0 \leq t_0$ such that

$$\int_0^t |f(s, p) - f(0, p)| ds < |t| \cdot \varepsilon \cdot \delta \quad (\text{B.4})$$

for every $|t| \leq \tilde{t}_0$ and each $p \in K_P$. Define now $A_p := \{s \in [0, t_0] : |f(s, p) - f(0, p)| > \delta\}$. From (B.4) we have

$$\delta \cdot \mu_L([0, t] \cap A_p) \leq \int_{[0, t] \cap A_p} |f(s, p) - f(0, p)| ds \leq \int_0^t |f(s, p) - f(0, p)| ds \leq |t| \cdot \varepsilon \cdot \delta,$$

hence $\mu_L([0, t] \cap A_p) \leq \varepsilon \cdot |t|$. Consequently, for $|t| < \tilde{t}_0$, we have

$$\begin{aligned} \int_0^t |G(f(s, p), s, p) - G(f(0, p), s, p)| ds &\leq 2 \cdot \sup_{K \times [0, t_0] \times K_P} |G| \int_{[0, t] \cap A_p} 1 ds + \int_{[0, t] \setminus A_p} \varepsilon ds \\ &\leq 2 \cdot \sup_{K \times [0, t_0] \times K_P} |G| \cdot \mu_L([0, t] \cap A_p) + \varepsilon |t| = \varepsilon |t| \left(2 \cdot \sup_{K \times [0, t_0] \times K_P} |G| + 1 \right). \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this proves (B.1).

To prove (B.3) we proceed similarly. Again we restrict our attention to $K \times [0, t_0] \times K_P$ and fix $\varepsilon > 0$. Let $\delta > 0$ be such that, for $x, y \in K$, $s \in [0, t_0]$ and $p, p' \in K_P$

$$|G(x, s, p) - G(y, s, p')| < \varepsilon$$

if $|x - y| < \delta$ and $|p - p'| < \delta$.

From the uniform regularity of $f(s, p)$, there exists a number $\tilde{\delta} > 0$ such that

$$\int_0^{t_0} |f(s, p) - f(s, p')| ds \leq \varepsilon \cdot \delta$$

if $p, p' \in K_P$ are such that $|p - p'| < \tilde{\delta}$. From that we deduce that the set $B_{pp'} := \{s \in [0, t_0] : |f(s, p) - f(s, p')| > \delta\}$ has measure smaller than ε if $|p - p'| < \tilde{\delta}$. Indeed, we can estimate

$$\delta \cdot \mu_L(B_{pp'}) \leq \int_{B_{pp'}} |f(s, p) - f(s, p')| ds \leq \int_0^{t_0} |f(s, p) - f(s, p')| ds \leq \varepsilon \cdot \delta.$$

Therefore for $|p - p'| \leq \min\{\delta, \tilde{\delta}\}$ we can estimate

$$\begin{aligned} \int_0^{t_0} |G(f(s, p), s, p) - G(f(s, p'), s, p')| ds &\leq \int_{B_{pp'}} |G(f(s, p), s, p) - G(f(s, p'), s, p')| ds \\ &\quad + \int_{[0, t_0] \setminus B_{pp'}} |G(f(s, p), s, p) - G(f(s, p'), s, p')| ds \\ &\leq 2 \cdot \sup_{K \times [0, t_0] \times K_P} |G| \cdot \mu_L(B_{pp'}) + \int_0^{t_0} \varepsilon ds \leq \varepsilon \left(2 \cdot \sup_{K \times [0, t_0] \times K_P} |G| + |t_0| \right). \end{aligned}$$

This proves (B.3). □

B.2 Ordinary differential equations

This section contains a revision of the theory of ordinary differential equations in a measurable setting. We formulate standard theorems about existence, uniqueness and regularity of solutions. We state these results after [Bressan & Piccoli, 2007] and give sketches of the proofs.

Carathéodory solutions

Consider an ordinary differential equation associated with a map $g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$,

$$\dot{x}(t) = g(x(t), t). \quad (\text{B.5})$$

By a (*Carathéodory*) *solution* of (B.5) on an interval $I = [t_0, t_1]$ we shall mean an AC map $t \mapsto x(t)$ which satisfies (B.5) a.e. For the solutions in the above sense one can develop the standard theory of existence, uniqueness, and parameter dependence, as done in [Bressan & Piccoli, 2007]. Let us recall the most important results of this theory. Assume the following:

$$t \mapsto g(x, t) \text{ is measurable for every } x, \text{ and } x \mapsto g(x, t) \text{ is continuous for every } t; \quad (\text{A})$$

$$g(x, t) \text{ is locally bounded and locally Lipschitz w.r.t. } x; \quad (\text{B})$$

that is, for every compact set $K \subset \mathbb{R}^n \times \mathbb{R}$ there exist constants C_K and L_K such that $|g(x, t)| \leq C_K$ and $|g(x, t) - g(y, t)| \leq L_K|x - y|$ for every $(x, t), (y, t) \in K$.

Theorem B.12 (existence and uniqueness of solutions). *Assuming that (A) and (B) hold, for every $x_0 \in \mathbb{R}^n$ there exists a unique solution $x(t, x_0)$ of (B.5) with the initial condition $x(t_0) = x_0$, defined on some interval $[t_0, t_0 + \varepsilon]$. If g is globally bounded and globally Lipschitz (so that the constants C_K and L_K in (B) can be chosen universally for all K 's), then the solution is also defined globally. Moreover, if $x(t, x_0)$ is defined on the interval $[t_0, t_1]$ then so are the solutions $x(t, x'_0)$ for x'_0 close enough to x_0 .*

Sketch of the proof. The proof uses the standard Picard's Method. One constructs a contracting map

$$A_{x_0} : x(t) \mapsto x_0 + \int_{t_0}^t g(x(\tau), \tau) d\tau$$

and uses it to define inductively a sequence of functions $x^0(t, x_0) = x_0$, $x^{n+1}(t, x_0) = A_{x_0}(x^n(t, x_0))$ which converges uniformly in t to the solution $x(t, x_0)$. The length of the interval $[t_0, t_1]$ on which the solution is well-defined depends on the Lipschitz bound of $g(x, t)$. The details can be found in [Bressan & Piccoli, 2007, Thm. 2.1.1]. \square

Parameter dependence

Consider now differential equation (B.5) with an additional parameter dependence

$$\dot{x}(t) = g(x(t), t, s), \quad (\text{B.6})$$

where $g : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}^n$. Assume the following:

$$t \mapsto g(x, t, s), \quad s \mapsto g(x, t, s) \text{ are measurable, and } x \mapsto g(x, t, s) \text{ is continuous;} \quad (A')$$

$$g(x, t, s) \text{ is locally bounded and locally Lipschitz w.r.t. } x; \quad (B')$$

that is, for every compact set $K \subset \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ there exist constants C_K, L_K such that $|g(x, t, s)| \leq C_K$ and $|g(x, t, s) - g(y, t, s)| \leq L_K|x - y|$ for every $(x, t, s), (y, t, s) \in K$.

Theorem B.13 (parameter dependence). *Assume that (A') and (B') hold, and denote by $x(t, x_0, s)$ the solution of (B.6) for a fixed parameter s and the initial condition $x(t_0, x_0, s) = x_0$ (we know that such solutions locally exist by Theorem B.12). Then the dependence $x_0 \mapsto x(t, x_0, s)$ is continuous, whereas, for any bounded measurable map $s \mapsto x_0(s)$, the map $s \mapsto x(t, x_0(s), s)$ is also bounded and measurable for every t .*

Sketch of the proof. As before one constructs a sequence $x^n(t, x_0, s)$ defined by means of the contracting map

$$A_{x_0, s} : x(t) \longmapsto x_0 + \int_{t_0}^t g(x(\tau), \tau, s) d\tau.$$

The sequence converges to the solution $x(t, x_0, s)$ uniformly w.r.t. t and x_0 , which implies continuity of the solution w.r.t. the initial value. If $x_0(s)$ is measurable w.r.t. s , so is the sequence $x^n(t, x_0(s), s)$. The limit $x(t, x_0(s), s)$ is measurable as a point-wise limit of measurable functions. Moreover, since $A_{x_0, s}$ is a contraction, $\|x(t, x_0(s), s)\|$ is bounded by a constant times $\|x_0(s)\|$. Details can be found in [Bressan & Piccoli, 2007]. \square

Assuming higher regularity of $g(x, t, s)$, one can prove a stronger result.

Theorem B.14 (differentiability w.r.t. the initial value). *Assume that the function $g(x, t, s)$ satisfies (A') and (B'), it is differentiable w.r.t. x , and the derivative $\frac{\partial g}{\partial x}(x, t, s)$ satisfies (A') and is locally bounded. Then the solution $x(t, x_0, s)$ of (B.6) is differentiable w.r.t. the initial condition x_0 . Moreover, the derivative $\frac{\partial x}{\partial x_0}(t, x_0, s)$ is continuous in x_0 , AC in t , and measurable in s .*

Sketch of the proof. We proceed again according to the standard method paying more attention to measurability. Consider a variation of (B.6)

$$\begin{aligned} \dot{x}(t) &= g(x(t), t, s), \\ \dot{y}(t) &= \frac{\partial g}{\partial x}(x(t), t, s) \end{aligned}$$

with the initial conditions $x(t_0) = x_0$ and $y(t_0) = \text{id}$. The above equations satisfy the assumptions of Theorem B.13, hence the solution $x(t, x_0, s)$ and $y(t, x_0, s)$ is a uniform

(in t and x_0) limit of the Picard's sequence $x^n(t, x_0, s)$ and $y^n(t, x_0, s)$. We observe that $\frac{\partial x^n}{\partial x_0}(t, x_0, s) = y^n(t, x_0, s)$, hence also $\frac{\partial x}{\partial x_0}(t, x_0, s) = y(t, x_0, s)$. The derivative $y(t, x_0, s)$ satisfies the regularity conditions by Theorem B.13.

Again a detailed proof (the only difference is the absence of the parameter s) can be found in [Bressan & Piccoli, 2007, Thm. 2.3.2]. \square

Gronwall Inequality

At the end of this section we will recall the following classical result.

Theorem B.15 (integral Gronwall Inequality). *Assume that $b : [0, T] \rightarrow \mathbb{R}$ is non-negative and integrable, and for almost every $t \in [0, T]$ we have*

$$b(t) \leq C \cdot \int_0^t b(s) ds,$$

where $C > 0$ is a constant. Then $b(t) = 0$ a.e.

For the proof see for instance [Evans, 2010, App. B].

B.3 E -homotopy type equations

In this section we consider linear PDEs of a special kind which are important in the notion of E -homotopy. We study their solutions in a weak sense and address a question about the existence of the trace.

E -homotopy type equations

Consider two functions $a(t, s)$ and $b(t, s)$ defined on a rectangle $K = [t_0, t_1] \times [0, 1] \ni (t, s)$. We will treat a and b as \mathbb{R} -valued, yet all results remain valid for \mathbb{R}^n -valued maps. Assume that a and b are bounded and measurable w.r.t. both variables separately. Let $c(t, s)$ be a fixed continuous function on K . We will say that the pair (a, b) is a *weak solution* (W -solution) of the differential equation

$$\partial_t b(t, s) = \partial_s a(t, s) + c(t, s)b(t, s)a(t, s) \quad (\text{B.7})$$

if for every function $\varphi \in C_0^\infty(K)$ the following equality holds:

$$- \iint_K \left[b(t, s) \partial_t \varphi(t, s) - a(t, s) \partial_s \varphi(t, s) + c(t, s)b(t, s)a(t, s)\varphi(t, s) \right] dt ds = 0. \quad (\text{B.8})$$

Observe that, since we assumed only measurability of a and b , the boundary values on ∂K are, in general, not well-defined.

Definition B.16. We say that a W-solution (a, b) of (B.7) has a *well-defined trace* if there exist bounded measurable maps $a_0, a_1 : [t_0, t_1] \rightarrow \mathbb{R}$ and $b_0, b_1 : [0, 1] \rightarrow \mathbb{R}$ such that, for every $\psi \in C^\infty(K)$, we have

$$\begin{aligned} & - \iint_K \left[b(t, s) \partial_t \psi(t, s) - a(t, s) \partial_s \psi(t, s) + c(t, s) b(t, s) a(t, s) \psi(t, s) \right] dt ds \\ & = \int_0^1 b_1(s) \psi(t_1, s) - b_0(s) \psi(t_0, s) ds + \int_{t_0}^{t_1} a_0(t) \psi(t, 0) - a_1(t) \psi(t, 1) dt. \end{aligned} \quad (\text{B.9})$$

In such a case we will call (a, b) a *WT-solution* of (B.7). The maps a_0, a_1 and b_0, b_1 will be called *traces* of a and b , respectively.

Remark B.17. Since the values of measurable functions are defined a.e. only, for a WT-solution (a, b) of (B.7) we will assume that the traces agree with the boundary values of a and b , i.e. $a_0(t) = a(t, 0)$, $a_1(t) = a(t, 1)$, $b_0(s) = b(t_0, s)$, and $b_1(t) = b(t_1, s)$.

Existence of the trace and properties of WT-solutions

Under certain regularity conditions, W-solutions of (B.7) are, in fact, WT-solutions.

Theorem B.18. *Let a and b be a bounded W-solution of (B.7). Assume in addition that a and b satisfy the following regularity conditions:*

$$\int_{t_0}^{t_1} \int_0^\varepsilon |a(t, s) - a(t, 0)| \frac{1}{\varepsilon} ds dt \xrightarrow{\varepsilon \rightarrow 0} 0, \quad \int_{t_0}^{t_1} \int_{1-\varepsilon}^1 |a(t, s) - a(t, 1)| \frac{1}{\varepsilon} ds dt \xrightarrow{\varepsilon \rightarrow 0} 0, \quad (\text{B.10})$$

$$\int_0^1 \int_{t_0}^{t_0+\varepsilon} |b(t, s) - b(t_0, s)| \frac{1}{\varepsilon} dt ds \xrightarrow{\varepsilon \rightarrow 0} 0, \quad \int_0^1 \int_{t_1-\varepsilon}^{t_1} |b(t, s) - b(t_1, s)| \frac{1}{\varepsilon} dt ds \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (\text{B.11})$$

Then (a, b) is a WT-solution of (B.7) and the traces $a(t, 0)$, $a(t, 1)$, $b(t_0, s)$, and $b(t_1, s)$ are well-defined.

Proof. Fix an element $\psi \in C^\infty(K)$ and choose $\varepsilon > 0$. The idea of the proof is standard: we will approximate ψ by another function $\varphi \in C_0^\infty(K)$ and, using (B.8) for φ and the regularity conditions, show that (B.9) holds with ε -accuracy.

Define a rectangle $K_\varepsilon := [t_0 + \varepsilon, t_1 - \varepsilon] \times [\varepsilon, 1 - \varepsilon] \subset K$, and choose a smooth "hat function" $\chi_{[a, b]} : [a, b] \rightarrow \mathbb{R}$ which satisfies the following conditions:

$$\chi_{[a, b]}(a) = 0 = \chi_{[a, b]}(b) = 0, \quad \chi_{[a, b]}(t) = 1 \text{ for } t \in [a + \varepsilon, b - \varepsilon] \quad \text{and} \quad \|D\chi_{[a, b]}\| < \frac{2}{\varepsilon}.$$

Now define $\chi(t, s) := \chi_{[t_0, t_1]}(t) \cdot \chi_{[0, 1]}(s)$. Obviously, $\chi \equiv 1$ on K_ε , $\chi \in C_0^\infty(K)$ and $\|D\chi_{[a, b]}\| < \frac{4}{\varepsilon}$. Moreover, $\partial_t \chi(t, s) = 0$ for $t \in [t_0 + \varepsilon, t_1 - \varepsilon]$, and $\partial_s \chi(t, s) = 0$ for $s \in [\varepsilon, 1 - \varepsilon]$.

Now define $\varphi := \chi \cdot \psi \in C_0^\infty(K)$ and $\tilde{\psi} := (1 - \chi)\psi$. Decomposing $\psi = \varphi + \tilde{\psi}$, we get

$$\begin{aligned} & - \iint_K [b\partial_t\psi - a\partial_s\psi + cba\psi] dt ds \\ &= - \iint_K [b\partial_t\varphi - a\partial_s\varphi + cba\varphi] dt ds - \iint_K [b\partial_t\tilde{\psi} - a\partial_s\tilde{\psi} + cba\tilde{\psi}] dt ds \stackrel{(B.8)}{=} \\ &= 0 - \iint_K b\partial_t\tilde{\psi} dt ds + \iint_K a\partial_s\tilde{\psi} dt ds - \iint_K cba\tilde{\psi} dt ds =: I_1 + I_2 + I_3. \end{aligned}$$

We will now concentrate on the tree last summands. Observe that

$$|I_3| = \left| \iint_K cba\tilde{\psi} dt ds \right| = \left| \iint_{K \setminus K_\varepsilon} cba\tilde{\psi} dt ds \right| \leq \mu(K \setminus K_\varepsilon) \|cba\tilde{\psi}\| \leq \varepsilon \cdot C_3,$$

where C_3 is a constant depending on $\|a\|$, $\|b\|$, $\|\psi\|$, and $\|c\|$.

Now

$$\begin{aligned} I_2 &= \iint_K a\partial_s\tilde{\psi} dt ds = \iint_{K \setminus K_\varepsilon} a\partial_s\tilde{\psi} dt ds \\ &= \iint_{K \setminus K_\varepsilon} a\psi\partial_s(1 - \chi) dt ds + \iint_{K \setminus K_\varepsilon} a(1 - \chi)\partial_s\psi dt ds. \end{aligned}$$

The last summand can be estimated by $\varepsilon \cdot C_2$ in the same way as I_3 (with C_2 depending additionally on $\|D\psi\|$). Now, since $\partial_s\chi(t, s) = 0$ for $s \in [\varepsilon, 1 - \varepsilon]$,

$$\iint_{K \setminus K_\varepsilon} a\psi\partial_s(1 - \chi) dt ds = - \int_{t_0}^{t_1} \int_0^\varepsilon a\psi\partial_s\chi ds dt - \int_{t_0}^{t_1} \int_{1-\varepsilon}^1 a\psi\partial_s\chi ds dt =: I_4 + I_5.$$

We can write I_4 as

$$\begin{aligned} I_4 &= \int_{t_0}^{t_1} [a(t, s)\psi(t, s) - a(t, 0)\psi(t, 0)] \partial_s\chi ds dt + \\ &+ \int_{t_0}^{t_1} a(t, 0)\psi(t, 0) \left[\int_0^\varepsilon \partial_s\chi ds \right] dt = I_6 + \int_{t_0}^{t_1} a(t, 0)\psi(t, 0) dt. \end{aligned}$$

Now

$$\begin{aligned} I_6 &= \int_{t_0}^{t_1} \int_0^\varepsilon (a(t, s) - a(t, 0)) \psi(t, 0) \partial_s\chi ds dt + \\ &+ \int_{t_0}^{t_1} \int_0^\varepsilon a(t, s) (\psi(t, s) - \psi(t, 0)) \partial_s\chi ds dt =: I_7 + I_8. \end{aligned}$$

Clearly,

$$|I_7| \leq \int_{t_0}^{t_1} \int_0^\varepsilon |a(t, s) - a(t, 0)| \cdot \|\psi\| \cdot \|D\chi\| ds dt \leq \int_{t_0}^{t_1} \int_0^\varepsilon |a(t, s) - a(t, 0)| \cdot \|\psi\| \frac{4}{\varepsilon} ds dt,$$

so by assumptions it converges to 0 as $\varepsilon \rightarrow 0$. Finally, using $|\psi(t, s) - \psi(t, 0)| \leq s \cdot \|D\psi\|$, we get

$$|I_8| \leq \int_{t_0}^{t_1} \int_0^\varepsilon \|a\| s \|D\psi\| \cdot \|D\chi\| ds dt \leq \int_{t_0}^{t_1} \int_0^\varepsilon \|a\| \varepsilon \|D\psi\| \frac{4}{\varepsilon} ds dt \leq \varepsilon \cdot C_8.$$

As a consequence, we get

$$I_4 \rightarrow \int_{t_0}^{t_1} a(t, 0) \psi(t, 0) dt \quad \text{as } \varepsilon \rightarrow 0.$$

Analogous estimations can be done for I_5 . As a result we get that

$$\left| I_2 - \int_{t_0}^{t_1} [a(t, 0) \psi(t, 0) - a(t, 1) \psi(t, 1)] dt \right| \xrightarrow{\varepsilon \rightarrow 0} 0.$$

We can repeat the above considerations for I_1 to prove that

$$\left| I_1 + \int_0^1 [b(t_0, s) \psi(t_0, s) - b(t_1, s) \psi(t_1, s)] ds \right| \xrightarrow{\varepsilon \rightarrow 0} 0.$$

The estimations for I_1 , I_2 and I_3 show that, for a fixed ψ , the equality (B.9) is satisfied with an accuracy converging to 0 as $\varepsilon \rightarrow 0$. \square

For WT-solutions we can formulate an uniqueness result.

Lemma B.19 (uniqueness of WT-solutions). *Let $a : K \rightarrow \mathbb{R}$ be a bounded measurable map (w.r.t. both variables separately), and let $b_0 : [0, 1] \rightarrow \mathbb{R}$ be any bounded measurable map. Then there exists at most one bounded measurable map $b : K \rightarrow \mathbb{R}$ such that (a, b) is a WT-solution of (B.7), and $b(t_0, s) = b_0(s)$. Moreover, the trace $b(t_1, s)$ is determined uniquely.*

Proof. Assume that $b(t, s)$ and $\tilde{b}(t, s)$ are two such solutions for a fixed $a(t, s)$. The difference $\delta b(t, s) := b(t, s) - \tilde{b}(t, s)$ is a bounded measurable map which is a WT-solution of the linear equation

$$\partial_t \delta b(t, s) = c(t, s) \delta b(t, s) a(t, s) \quad (\text{B.12})$$

such that $\delta b(t_0, s) = 0$. Let us define $B(\tau, s) := 0 + \int_{t_0}^\tau c(t, s) \delta b(t, s) a(t, s) dt$. Clearly, $B(t, s)$ is ACB w.r.t. t and measurable w.r.t. s . Moreover, we have $\partial_t B(t, s) = c(t, s) \delta b(t, s) a(t, s)$ in the sense of Carathéodory and, since $B(t, s)$ is continuous w.r.t. t , also WT. Consequently, $(B - \delta b)$ satisfies $\partial_t (B - \delta b) \stackrel{\text{WT}}{=} 0$ and, since $B(t_0, s) = \delta b(t_0, s) = 0$, we have

$$- \iint_K (B - \delta b)(t, s) \partial_t \psi(t, s) dt ds = \int_0^1 (B - \delta b)(t_1, s) \psi(t_1, s) ds \quad (\text{B.13})$$

for every $\psi \in C^\infty(K)$. Taking $\psi(t, s) = \phi(s)$, where $\phi \in C^\infty(I)$, we get that $\int_0^1 (B - \delta b)(t_1, s) \phi(s) ds = 0$, thus $B(t_1, s) = \delta b(t_1, s)$ a.e. In the light of this observation (B.13) reads as

$$\iint_K (B - \delta b)(t, s) \partial_t \psi(t, s) dt ds = 0,$$

for every $\psi \in C^\infty(K)$. Since $\partial_t \psi$ can be an arbitrary smooth function, we conclude that $B(t, s) = \delta b(t, s)$ a.e. Consequently, δb is a Carathéodory solution of (B.12). Now observe that

$$|\delta b(\tau, s)| = \left| \int_{t_0}^{\tau} \partial_t \delta b(t, s) dt \right| \leq \int_{t_0}^{\tau} |\partial_t \delta b(t, s)| dt = \int_{t_0}^{\tau} |c \cdot a| |\delta b(t, s)| dt,$$

which, in view of the integral Gronwall Inequality B.15, implies $\delta b = 0$. \square

Control theory

In this part we recall basic definitions from control theory. Later we formulate the Pontryagin maximum principle in its classical form.

Definition C.1. A *control system* on a manifold M is a map

$$f : M \times U \longrightarrow TM, \quad (\text{C.1})$$

such that, for every fixed $u \in U$, the map $f(\cdot, u) : M \longrightarrow TM$ is a C^1 -vector field. We assume that U is a subset of some Euclidean space \mathbb{R}^r and that the maps $f : M \times U \longrightarrow TM$ and $T_x f : TM \times U \longrightarrow TTM$ are continuous.

Choose now an *admissible control*, i.e., a bounded measurable function $u : [t_0, t_1] \rightarrow U$. We can consider a time-dependent differential equation on M

$$\dot{x}(t) = f(x(t), u(t)),$$

with a fixed initial condition $x(t_0) = x_0$. The solution $x(t)$ of the above is called a *trajectory* of a control system (C.1) associated with the control $u(t)$, and the pair $(x(t), u(t))$ is called a *controlled pair*.

Let us now introduce the *total cost* of the controlled pair $(x(t), u(t))$

$$\mathcal{J}(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} L(x(t), u(t)) dt,$$

where on the integrand $L : M \times U \longrightarrow \mathbb{R}$ (the *cost function*) we put the same regularity assumptions as on f , namely, $L : M \times U \longrightarrow \mathbb{R}$ and $T_x L : TM \times U \longrightarrow \mathbb{R}$ are continuous maps.

Given two points x_0 and x_1 we can introduce an *optimal control problem*:

$$\begin{aligned} & \text{minimise the total cost } \mathcal{J}(x(\cdot), u(\cdot)) \text{ over all controlled pairs } (x(t), u(t)) \\ & \text{(with all possible time intervals } t \in [t_0, t_1]) \text{ s.t. } x(t_0) = x_0 \text{ and } x(t_1) = x_1. \end{aligned} \quad (\tilde{\text{P}})$$

Let now $\iota_0 : S_0 \hookrightarrow M$ and $\iota_1 : S_1 \hookrightarrow M$ be two immersed submanifolds of M . We define the following OCP with *general boundary conditions*:

$$\begin{aligned} & \text{minimise the total cost } \mathcal{J}(x(\cdot), u(\cdot)) \text{ over all controlled pairs } (x(t), u(t)) \\ & \text{(with all possible time intervals } t \in [t_0, t_1]) \text{ such that } x(t_0) \in S_0 \text{ and } x(t_1) \in S_1. \end{aligned} \quad (\hat{\text{P}})$$

Necessary optimality conditions for the problem $(\tilde{\text{P}})$ are the following.

Theorem C.2 (the PMP). *Let $(x(t), u(t))$, with $t \in [t_0, t_1]$, be a controlled pair of (C.1) solving the optimal control problem (\hat{P}) . Then there exists a curve $p : [t_0, t_1] \rightarrow T^*M$ covering $x(t)$ and a constant $p_0 \leq 0$ such that the following holds:*

- *the curve $p(t)$ is a trajectory of the time-dependent family of Hamiltonian vector fields \mathcal{X}_{H_t} for the canonical symplectic structure on T^*M and Hamiltonians $H_t(x, p) := H(x, p, u(t))$, where*

$$H(x, p, u) = \langle f(x, u), p \rangle + p_0 L(x, u);$$

- *the control u satisfies the “maximum principle”*

$$H(x(t), p(t), u(t)) = \sup_{v \in U} H(x(t), p(t), v)$$

and $H(x(t), p(t), u(t)) = 0$ at every regular point t of u ;

- *if $p_0 = 0$ the covector $p(t)$ is nowhere-vanishing.*

For the problem (\hat{P}) we have more specific conditions.

Theorem C.3 (the PMP for general boundary conditions). *Let $(x(t), u(t))$, with $t \in [t_0, t_1]$, be a controlled pair of (C.1) solving the optimal control problem (\hat{P}) . Then there exists a curve $p : [t_0, t_1] \rightarrow T^*M$ covering $x(t)$ and a constant $p_0 \leq 0$ which satisfy the assertion of Theorem C.2 and, additionally, $p(t)$ satisfies the following transversality conditions: $p(t_0)$ annihilates $T_{x(t_0)}S_0$ and $p(t_1)$ annihilates $T_{x(t_1)}S_1$.*

For the original proof of the above theorems we refer the reader to [Pontryagin *et al.*, 1962]. Recent references are [Agrachev & Sachkov, 2004] and [Barbero-Liñán & Muñoz-Lecanda, 2009].

Geometry and Topology

D.1 Separation of convex cones

Geometrically, Pontryagin maximum principle describes the separation of certain cones associated with the optimal control problem, which live in the fibres of the algebroid $E \oplus T\mathbb{R}$, along an optimal trajectory. Therefore we need some technical results concerning the separation of convex cones.

Definition D.1. Two convex sets K_1 and K_2 in a vector space V are *separable* iff there exists a non-zero covector $\varphi \in V^*$ such that

$$\langle k_1, \varphi \rangle \geq \langle k_2, \varphi \rangle \quad \text{for every } k_1 \in K_1 \text{ and } k_2 \in K_2. \quad (\text{D.1})$$

We say that K_1 and K_2 are *strictly separable* iff there exists a non-zero covector $\varphi \in V^*$ and numbers $a, b \in \mathbb{R}$ such that

$$\langle k_1, \varphi \rangle \geq a > b \geq \langle k_2, \varphi \rangle \quad \text{for every } k_1 \in K_1 \text{ and } k_2 \in K_2.$$

A basic fact from the theory of convex sets in a finite dimensional space is the following

Theorem D.2 (separation). *Two disjoint convex sets in a finite dimensional vector space are separable. If, in addition, these sets are closed, and one of them is compact, they are strictly separable.*

For the proof see [Giannessi, 2005]. We will also need the following fact.

Lemma D.3. *The convex sets K_1 and K_2 in a finite-dimensional space V are separable if and only if $\text{cl}(K_1)$ and $\text{cl}(K_2)$ are separable.*

Proof. If (D.1) holds for every $k_1 \in K_1$ and $k_2 \in K_2$ then, since the weak inequality is preserved under taking limits, it also holds for every $k_1 \in \text{cl}(K_1)$ and $k_2 \in \text{cl}(K_2)$.

Conversely, if (D.1) holds for every $k_1 \in \text{cl}(K_1)$ and $k_2 \in \text{cl}(K_2)$, it is also true on smaller sets $K_1 \subset \text{cl}(K_1)$ and $K_2 \subset \text{cl}(K_2)$. \square

Definition D.4. By a *cone* in a vector space V we will mean a set K which is invariant under positive homotheties, i.e.,

$$t \cdot k \in K \quad \text{whenever } k \in K \text{ and } t > 0.$$

Remark D.5. If two convex sets K_1 and K_2 in V are separable, and one of them, say K_1 , is a cone, then the separating covector φ satisfies

$$\langle k_1, \varphi \rangle \geq 0 \geq \langle k_2, \varphi \rangle \quad \text{for every } k_1 \in K_1 \text{ and } k_2 \in K_2. \quad (\text{D.2})$$

Indeed, since K_1 is invariant under homotheties, the image $\varphi(K_1)$ contains numbers arbitrary close to 0, hence from (D.1) it satisfies $0 \geq \langle k_2, \varphi \rangle$ for all $k_2 \in K_2$. On the other hand, if $\langle k_1, \varphi \rangle < 0$ for some $k_1 \in K_1$, then the image $\langle K_1, \varphi \rangle$ would contain arbitrarily big negative numbers, and hence (D.1) would not hold.

Above observation has two simple but important consequences. First of all, the separating covector φ vanishes on the intersection $K_1 \cap K_2$. Secondly, if one of the sets K_i contains an affine subspace $k_i + S$, where $S \subset V$ is a linear subspace, then φ vanishes on S .

Now we prove a geometric characterisation of non-separability in a certain geometric setting.

Lemma D.6. Consider a vector space $V = W \oplus \mathbb{R}$, denote by Λ a ray in \mathbb{R} spanned by a vector λ , i.e., $\Lambda = \mathbb{R}_+ \cdot \lambda \subset \mathbb{R}$, and let $S \subset W$ be a linear subspace. Let K_1 be a convex cone in V , and denote by K_2 the convex cone $S \oplus \Lambda$. The cones K_1 and K_2 are not separable iff there exists a vector $k \in K_1 \cap K_2$ and vectors $e_1, e_2, \dots, e_m \in W$ such that

$$\text{span}\{e_1, e_2, \dots, e_m, S\} = W, \quad (\text{D.3})$$

$$\text{vectors } k \pm e_1, \dots, k \pm e_m \text{ belong to } K_1, \quad (\text{D.4})$$

where we identify $e_i \in W$ with $e_i + \theta \in W \oplus \mathbb{R} = V$.

Proof. Assume that vectors k, e_1, \dots, e_m satisfy conditions (D.3) and (D.4). If φ is a non-zero covector separating K_1 and K_2 then, due to Remark D.5, φ vanishes on k and S . Moreover, $0 \leq \langle k \pm e_i, \varphi \rangle = \pm \langle e_i, \varphi \rangle$, hence $\langle e_i, \varphi \rangle = 0$. Since $\text{span}\{k, e_1, \dots, e_m, S\} = V$, the covector φ is null, which gives a contradiction.

The opposite implication is harder to prove. We will make an inductive argument with respect to $\dim S$.

Assume that $\dim S = 0$; i.e., $K_2 = 0 \oplus \Lambda = \mathbb{R}_+ \cdot \lambda$. If K_1 and K_2 are not separable then

$$\text{span}\{K_1, \lambda\} = V \quad \text{and} \quad \lambda \in K_1.$$

Indeed, if $V' = \text{span}\{K_1, \lambda\} \subsetneq V$, then a non-zero covector vanishing on V' will separate K_1 and K_2 . Secondly, if $\lambda \notin K_1$ then, by Theorem D.2, convex sets K_1 and $\{\lambda\}$ can be separated by a covector φ . Since K_1 is a cone we have $\langle k_1, \varphi \rangle \geq 0 \geq \langle \lambda, \varphi \rangle$ for $k_1 \in K_1$ (compare Remark D.5), hence φ separates also K_1 and $\mathbb{R}_+ \cdot \lambda = K_2$.

We deduce that $\lambda \in K_1$, and that there exist vectors $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_m \in K_2$ such that $\{\lambda, \bar{e}_1, \dots, \bar{e}_m\}$ is a basis of V . Obviously each \bar{e}_i is of the form $\bar{e}_i = a_i \cdot \lambda + \hat{e}_i$, where a_i are numbers and $\{\hat{e}_1, \dots, \hat{e}_m\}$ is a basis of W .

Now fix i and consider a vector $\lambda - \frac{1}{N}\hat{e}_i$, where $N \in \mathbb{N}$. If $\lambda - \frac{1}{N}\hat{e}_i \notin K_1$ for all $N \in \mathbb{N}$, then, by Theorem D.2 and Remark D.5, there exist covectors φ_N such that

$$\langle k_1, \varphi_N \rangle \geq 0 \quad \text{for all } k_1 \in K_1 \quad \text{and} \quad \langle \lambda - \frac{1}{N}\hat{e}_i, \varphi_N \rangle \leq 0.$$

We may assume that all φ_N are normalised to 1 and choose a subsequence converging to $\varphi_0 \in V^*$. Clearly, φ_0 is non-zero (it is normalised to 1) and

$$\langle k_1, \varphi_0 \rangle \geq 0 \quad \text{for all } k_1 \in K_1 \quad \text{and} \quad \langle \lambda, \varphi_0 \rangle \leq 0,$$

that is, φ_0 separates K_1 and $K_2 = \mathbb{R}_+ \cdot \lambda$ against the assumptions.

To sum up, we proved that K_1 contains elements λ , $a_i \lambda + \widehat{e}_i$ and $\lambda - \frac{1}{N_i} \widehat{e}_i$. It is clear that some convex combinations of these vectors, after rescaling, are of the form $\lambda + e_i$ and $\lambda - e_i$, where e_i is parallel to \widehat{e}_i .

Now consider $\dim S > 0$. We can split $S = S' \oplus \widehat{S}$, where $S' = \mathbb{R} \cdot s$ is one-dimensional. By Remark D.5, if K_1 and K_2 are separable, then the separating covector φ vanishes on S' . It follows that K_1 and K_2 are separable in V if and only if $K'_1 := K_1/S'$ and $K'_2 := K_2/S' = \widehat{S} \oplus \Lambda$ are separable in $V' := V/S' = W/S' \oplus \mathbb{R}$. By the inductive assumption there exists vectors $k \in K_1/S' \cap K_2/S'$ and $\widehat{e}_1, \dots, \widehat{e}_m \in W/S'$ such that conditions (D.3) and (D.4) are satisfied for $W' = W/S'$ and K'_1 . In other words, there exists a vector $k \in S \oplus \Lambda$, vectors $\widehat{e}_1, \dots, \widehat{e}_m \in W$, and numbers a, a_i, b_i such that vectors $k + a \cdot s$, $k + a_i \cdot s + \widehat{e}_i$ and $k + b_i \cdot s - \widehat{e}_i$ belong to K_1 and $\text{span}\{\widehat{e}_1, \dots, \widehat{e}_m, S\} = \text{span}\{\widehat{e}_1, \dots, \widehat{e}_m, \widehat{S}, s\} = \text{span}\{W/S', s\} = W$.

From convexity of K_1 we deduce that $k + \frac{a_i + b_i}{2} \cdot s \in K_1$. Now, either all numbers $\frac{a_i + b_i}{2}$ are equal a , and then vectors $k + a \cdot s$ and $e_i = \widehat{e}_i + (a_i - a) \cdot s$ satisfy the assertion, or the interval $k + [c, d] \cdot s$, where $a, \frac{a_i + b_i}{2} \in [c, d]$, is entirely contained in K_1 . In the second case it is quite clear that some convex combinations of vectors $k + c \cdot s$, $k + d \cdot s$, $k + a_i \cdot s + \widehat{e}_i$ and $k + b_i \cdot s - \widehat{e}_i$ are of the form $k + \widetilde{a}_i \cdot s + \widetilde{e}_i$ and $k + \widetilde{b}_i \cdot s - \widetilde{e}_i$, where $\frac{\widetilde{a}_i + \widetilde{b}_i}{2} = \frac{c + d}{2} = \widetilde{a}$ and \widetilde{e}_i is parallel to \widehat{e}_i . Therefore we are in the first case again.

The inductive argument is now complete. \square

D.2 Simple topological lemmas

Lemma D.7. *Every continuous map $\Psi : B^m(0, 1) \rightarrow \mathbb{R}^m$ which satisfies the inequality*

$$\|\Psi(\vec{r}) - \vec{r}\| \leq \frac{1}{2} \quad \text{for all } \vec{r} \in B^m(0, 1),$$

contains point $0 \in \mathbb{R}^m$ in its image.

Proof. The map

$$H(\vec{R}, t) := (1 - t)\vec{R} + t\Psi(\vec{R}),$$

where $t \in [0, 1]$ and $\vec{R} \in \partial B^m(0, 1)$ is a homotopy between $\partial B^m(0, 1)$ and $\Psi(\partial B^m(0, 1))$. Observe that $\|H(\vec{R}, t) - \vec{R}\| = t\|\Psi(\vec{R}) - \vec{R}\| \leq \frac{1}{2}$, hence

$$\|H(\vec{R}, t)\| \geq \|\vec{R}\| - \|H(\vec{R}, t) - \vec{R}\| \geq 1 - \frac{1}{2} \geq \frac{1}{2},$$

and consequently H takes values in $\mathbb{R}^m \setminus \{0\}$.

Assume that $0 \notin \text{Im } \Psi$. Then the map

$$S(\vec{R}, t) := \Psi((1 - t)\vec{R})$$

is a homotopy between $\Psi(\partial B^m(0, 1))$ and $\Psi(0)$, which takes values in $\text{Im } \Phi \subset \mathbb{R}^m \setminus \{0\}$.

The composition

$$H \circ S(\vec{R}, t) = \begin{cases} H(\vec{R}, 2t) & \text{for } t < \frac{1}{2}, \\ S(\vec{R}, 2t - 1) & \text{for } t \geq \frac{1}{2} \end{cases}$$

is a contraction of $\partial B^m(0, 1)$ to the point $\Psi(0)$, which takes values in $\mathbb{R}^m \setminus \{0\}$. On the other hand, the sphere $\partial B^m(0, 1)$ is not contractible in $\mathbb{R}^m \setminus \{0\}$. \square

Lemma D.8. *Let $\Psi_1 : B^{m_1}(0, 1) \oplus \theta_{m_2} \rightarrow \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2}$ and $\Psi_2 : \theta_{m_1} \oplus B^{m_2}(0, 1) \rightarrow \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2}$ be two continuous maps satisfying, for a fixed vector $\vec{k}_0 \in \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2}$, the following inequalities:*

$$\begin{aligned} \|\Psi_1(\vec{r}) - (\vec{r} + \vec{k}_0)\| &\leq \frac{1}{4} \quad \text{for } \vec{r} \in B^{m_1}(0, 1) \oplus \theta_{m_2} \text{ and} \\ \|\Psi_2(\vec{s}) - (\vec{s} + \vec{k}_0)\| &\leq \frac{1}{4} \quad \text{for } \vec{s} \in \theta_{m_1} \oplus B^{m_2}(0, 1). \end{aligned}$$

Then the images $\text{Im } \Psi_1$ and $\text{Im } \Psi_2$ have a non-empty intersection.

Proof. Consider a continuous map

$$B^{m_1+m_2}(0, 1) = B^{m_1}(0, 1) \oplus B^{m_2}(0, 1) \ni (\vec{r} + \vec{s}) \xrightarrow{\Psi} \Psi_1(\vec{r}) - \Psi_2(-\vec{s}) \in \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2}.$$

Now

$$\begin{aligned} \|\Psi(\vec{r}, \vec{s}) - (\vec{r} + \vec{s})\| &= \left\| \Psi_1(\vec{r}) - (\vec{r} + \vec{k}_0) - \left(\Psi_2(-\vec{s}) - (-\vec{s} + \vec{k}_0) \right) \right\| \leq \\ &\leq \left\| \Psi_1(\vec{r}) - (\vec{r} + \vec{k}_0) \right\| + \left\| \Psi_2(-\vec{s}) - (-\vec{s} + \vec{k}_0) \right\| \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \end{aligned}$$

By Lemma D.7 point 0 lies in the image of Ψ . Consequently, $\Psi_1(\vec{r}_0) = \Psi_2(-\vec{s}_0)$ for some \vec{r}_0 and \vec{s}_0 . \square

Bibliography

- [Agrachev & Gamkrelidze, 2006] Agrachev, A.A., & Gamkrelidze, R.V. 2006. The Pontryagin Maximum Principle 50 years later. *Proceedings of the Steklov Institute of Mathematics*, **253**, 4–12.
- [Agrachev & Sachkov, 2004] Agrachev, A.A., & Sachkov, Y.L. 2004. *Control theory from the geometric viewpoint*. Encyclopaedia of mathematical sciences, no. 2. Springer.
- [Agrachev & Sarychev, 1996] Agrachev, A.A., & Sarychev, A.V. 1996. Abnormal sub-Riemannian geodesics: Morse index and rigidity. *Annales de l’Institut Henri Poincaré. Analyse non linéaire*, **13**(6), 635–690.
- [Agrachev & Sarychev, 1998] Agrachev, A.A., & Sarychev, A.V. 1998. On abnormal extremals for Lagrange variational problems. *Journal of Mathematical Systems Estimation and Control*, **8**, 87–118.
- [Almeida, 1980] Almeida, R. 1980. *Teoris de Lie para os groupoides diferenciaries*. Ph.D. thesis, Sao Paulo.
- [Almeida & Kumpera, 1981] Almeida, R., & Kumpera, A. 1981. Structure produit dans la catégorie des algébroïdes de Lie. *Anais da Academia Brasileira de Ciências*, **53**, 247–250.
- [Almeida & Molino, 1985] Almeida, R., & Molino, P. 1985. Suites d’Atiyah et feuilletages transversalement complets. *Comptes Rendus de l’Académie des Sciences, Série A*, **300**, 13–15.
- [Barbero-Liñán & Muñoz-Lecanda, 2009] Barbero-Liñán, M., & Muñoz-Lecanda, M.C. 2009. Geometric approach to Pontryagin’s maximum principle. *Acta applicandae mathematicae*, **108**(2), 429–485.
- [Bloch, 2003] Bloch, A. 2003. *Nonholonomic mechanics and control*. Interdisciplinary applied mathematics: Systems and control. Springer.
- [Bonnard & Trélat, 2001] Bonnard, B., & Trélat, E. 2001. On the role of abnormal minimizers in sub-Riemannian geometry. *Annales de la Faculté des Sciences de Toulouse Mathématiques*, **10**(3), 405–491.
- [Bressan & Piccoli, 2007] Bressan, A., & Piccoli, B. 2007. *Introduction to the mathematical theory of control*. AIMS series on applied mathematics. American Institute of Mathematical Sciences.

- [Cattaneo & Felder, 2004] Cattaneo, A.S., & Felder, G. 2004. Coisotropic submanifolds in Poisson geometry and branes in the Poisson sigma model. *Letters in Mathematical Physics*, **69**(1), 157–175.
- [Cendra *et al.*, 1998] Cendra, H., Holm, D.D., Marsden, J.E., & Ratiu, T.S. 1998. Lagrangian reduction, the Euler-Poincaré equations, and semidirect products. *Pages 1–25 of: Geometry of Differential Equations*. American Mathematical Society Translations, vol. 186. American Mathematical Society.
- [Chaplygin, 1911] Chaplygin, S.A. 1911. On the theory of the motion of nonholonomic systems. Theorem on the reducing multiplier. *Mat. Sbornik*, **28**(2), 303–314.
- [Clarke, 1976] Clarke, F. 1976. The maximum principle under minimal hypotheses. *SIAM Journal on Control and Optimization*, **14**, 1078.
- [Clarke, 2005a] Clarke, F. 2005a. The maximum principle in optimal control, then and now. *Control and Cybernetics*, **34**(3), 709.
- [Clarke, 2005b] Clarke, F. 2005b. *Necessary conditions in dynamic optimization*. Memoirs of the American Mathematical Society, no. 816. American Mathematical Society.
- [Cortés & Martínez, 2004] Cortés, J., & Martínez, E. 2004. Mechanical control systems on Lie algebroids. *IMA Journal of Mathematical Control and Information*, **21**(4), 457.
- [Cortés *et al.*, 2006] Cortés, J., De Leon, M., Marrero, J.C., De Diego, D.M., & Martinez, E. 2006. A survey of Lagrangian mechanics and control on Lie algebroids and groupoids. *International Journal of Geometric Methods in Modern Physics*, **3**(3), 509–558.
- [Coste *et al.*, 1987] Coste, A., Dazord, P., & Weinstein, A. 1987. Groupoïdes symplectiques. *Pages 1–62 of: Publications Département de Mathématiques*. Université Claude Bernard-Lyon I.
- [Courant, 1990] Courant, T.J. 1990. Dirac manifolds. *Transactions of the American Mathematical Society*, **319**(2), 631–661.
- [Crainic & Fernandes, 2003] Crainic, M., & Fernandes, R.L. 2003. Integrability of Lie brackets. *The Annals of Mathematics*, **157**(2), 575–620.
- [Crainic & Fernandes, 2004] Crainic, M., & Fernandes, R.L. 2004. Integrability of Poisson brackets. *Journal of Differential Geometry*, **66**(1), 71–137.
- [Duistermaat & Kolk, 2000] Duistermaat, J.J., & Kolk, J.A.C. 2000. *Lie groups*. Universitext. Springer.
- [Evans, 2010] Evans, L.C. 2010. *Partial differential equations*. Graduate studies in mathematics. American Mathematical Society.
- [Gamkrelidze *et al.*, 1978] Gamkrelidze, R.V., Makowski, K., & Berkovitz, L. 1978. *Principles of optimal control theory*. Plenum Press.

- [Giannessi, 2005] Giannessi, F. 2005. *Separation of sets and optimality conditions*. Constrained optimization and image space analysis. Springer.
- [Grabowska & Grabowski, 2008] Grabowska, K., & Grabowski, J. 2008. Variational calculus with constraints on general algebroids. *Journal of Physics A: Mathematical and Theoretical*, **41**, 175204.
- [Grabowska *et al.*, 2006] Grabowska, K., Grabowski, J., & Urbański, P. 2006. Geometrical mechanics on algebroids. *International Journal of Geometric Methods in Modern Physics*, **3**(3), 559–575.
- [Grabowski & Józwickowski, 2011] Grabowski, J., & Józwickowski, M. 2011. Pontryagin maximum principle on almost Lie algebroids. *SIAM Journal on Control and Optimization*, **49**(3), 1306–1357.
- [Grabowski & Rotkiewicz, 2009] Grabowski, J., & Rotkiewicz, M. 2009. Higher vector bundles and multi-graded symplectic manifolds. *Journal of Geometry and Physics*, **59**(9), 1285–1305.
- [Grabowski & Urbański, 1997] Grabowski, J., & Urbański, P. 1997. Lie algebroids and Poisson–Nijenhuis structures. *Reports on Mathematical Physics*, **40**(2), 195–208.
- [Grabowski & Urbański, 1999] Grabowski, J., & Urbański, P. 1999. Algebroids — general differential calculi on vector bundles. *Journal of Geometry and Physics*, **31**(2-3), 111–141.
- [Grabowski *et al.*, 2009] Grabowski, J., De Leon, M., Marrero, J.C., & De Diego, D.M. 2009. Nonholonomic constraints: A new viewpoint. *Journal of Mathematical Physics*, **50**, 013520.
- [Jurdjevic, 1997] Jurdjevic, V. 1997. *Geometric control theory*. Cambridge studies in advanced mathematics. Cambridge University Press.
- [Konieczna & Urbański, 1999] Konieczna, K., & Urbański, P. 1999. Double vector bundles and duality. *Archivum Mathematicum*, **35**(1), 59–95.
- [Kubarski, 1994] Kubarski, J. 1994. Invariant cohomology of regular Lie algebroids. *Pages 26–30 of: Proceedings of the VIIth International Colloquium on Differential Geometry, (July 1994, Spain)*.
- [Lang, 1985] Lang, S. 1985. *Differential manifolds*. Springer-Verlag.
- [Langerock, 2003a] Langerock, B. 2003a. *Generalised connections and applications to control theory*. Ph.D. thesis, Ghent University.
- [Langerock, 2003b] Langerock, B. 2003b. Geometric aspects of the maximum principle and lifts over a bundle map. *Acta Applicandae Mathematicae*, **77**(1), 71–104.

- [León *et al.*, 2005] León, M., Marrero, J.C., & Martínez, E. 2005. Lagrangian submanifolds and dynamics on Lie algebroids. *Journal of Physics A: Mathematical and General*, **38**, R241.
- [Libermann, 1996] Libermann, P. 1996. Lie algebroids and mechanics. *Arch. Math. (Brno)*, **32**, 147–162.
- [Łojasiewicz & Ferreira, 1988] Łojasiewicz, S., & Ferreira, A.V. 1988. *An introduction to the theory of real functions*. A Wiley interscience publication. John Wiley & Sons Inc.
- [Mackenzie, 1987] Mackenzie, K. 1987. *Lie groupoids and Lie algebroids in differential geometry*. London Mathematical Society lecture note series. Cambridge University Press.
- [Mackenzie, 2005] Mackenzie, K. 2005. *General theory of lie groupoids and lie algebroids*. London Mathematical Society lecture note series. Cambridge University Press.
- [Mackenzie & Xu, 2000] Mackenzie, K.C.H., & Xu, P. 2000. Integration of Lie bialgebroids. *Topology*, **39**(3), 445–467.
- [Martínez, 2001a] Martínez, E. 2001a. Geometric formulation of mechanics on Lie algebroids. *Pages 209–222 of: Proceedings of the VIII Fall Workshop on Geometry and Physics (1999, Medina del Campo), Publicaciones de la RSME*, vol. 2.
- [Martínez, 2001b] Martínez, E. 2001b. Lagrangian mechanics on Lie algebroids. *Acta Applicandae Mathematicae*, **67**(3), 295–320.
- [Martínez, 2004] Martínez, E. 2004. Reduction in optimal control theory. *Reports on Mathematical Physics*, **53**(1), 79–90.
- [Martínez, 2005] Martínez, E. 2005. Classical field theory on Lie algebroids: variational aspects. *Journal of Physics A: Mathematical and General*, **38**, 7145.
- [Martínez, 2007] Martínez, E. 2007. Lie Algebroids in Classical Mechanics and Optimal Control. *Symmetry, Integrability and Geometry: Methods and Applications*, **3**, 050.
- [Martínez, 2008] Martínez, E. 2008. Variational calculus on Lie algebroids. *ESAIM: Control, Optimisation and Calculus of Variations*, **14**(2), 356–380.
- [Moerdijk & Crainic, 2001] Moerdijk, I., & Crainic, M. 2001. Foliation groupoids and their cyclic homology. *Advances in Mathematics*, **157**(2), 177–197.
- [Moerdijk & Mrčun, 2003] Moerdijk, I., & Mrčun, J. 2003. *Introduction to foliations and Lie groupoids*. Cambridge Studies in Advanced Mathematics. Cambridge University Press.
- [Montgomery, 1990] Montgomery, R. 1990. Isoholonomic problems and some applications. *Communications in Mathematical Physics*, **128**(3), 565–592.

- [Montgomery, 1994] Montgomery, R. 1994. Abnormal minimizers. *SIAM Journal on control and optimization*, **32**(6), 1605–1620.
- [Neimark & Fufaev, 1972] Neimark, J.I., & Fufaev, N.A. 1972. *Dynamics of Nonholonomic Systems*. Translations of mathematical monographs. American Mathematical Society.
- [Pontryagin *et al.*, 1962] Pontryagin, L. S., Boltjanskij, V. G., Gamkrelidze, R. V., & Miscenko, E. F. 1962. *The mathematical theory of optimal processes*. Interscience Publishers.
- [Pradines, 1966] Pradines, J. 1966. Théorie de Lie pour les groupoides différentiable. *Comptes Rendus de l'Académie des Sciences, Série A*, **263**, 907–910.
- [Pradines, 1967a] Pradines, J. 1967a. Géométrie différentielle au-dessus d'un groupoïde. *Comptes Rendus de l'Académie des Sciences, Série A*, **266**, 1194–1196.
- [Pradines, 1967b] Pradines, J. 1967b. Théorie de Lie pour les groupoïdes différentiables. Calcul différentiel dans la catégorie des groupoides infinitésimaux. *Comptes Rendus de l'Académie des Sciences, Série A*, **264**, 245–248.
- [Pradines, 1968] Pradines, J. 1968. Troisième théorème de Lie pour les groupoïdes différentiables. *Comptes Rendus de l'Académie des Sciences, Série A*, **267**, 21–23.
- [Silva & Weinstein, 1999] Silva, A.C., & Weinstein, A. 1999. *Geometric models for non-commutative algebras*. Berkeley mathematics lecture notes. American Mathematical Society.
- [Tulczyjew, 1974] Tulczyjew, W.M. 1974. Hamiltonian systems, Lagrangian systems and the Legendre transformation. *Symposia Mathematica*, **14**, 247–258.
- [Tulczyjew & Urbanski, 1999] Tulczyjew, W.M., & Urbanski, P. 1999. A slow and careful Legendre transformation for singular Lagrangians, The Infeld Centennial Meeting (Warsaw, 1998). *Acta Physica Polonica B*, **30**, 2909–2978.
- [Weinstein, 1987] Weinstein, A. 1987. Symplectic groupoids and Poisson manifolds. *Bulletin of the American Mathematical Society*, **16**(1), 101–104.
- [Weinstein, 1988] Weinstein, A. 1988. Coisotropic calculus and Poisson groupoids. *Journal of the Mathematical Society of Japan*, **40**(4), 705–727.
- [Weinstein, 1996] Weinstein, A. 1996. Lagrangian mechanics and groupoids. *Fields Institute Communications*, **7**, 207–232.
- [Weinstein & Xu, 1991] Weinstein, A., & Xu, P. 1991. Extensions of symplectic groupoids and quantization. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, **417**, 159–190.

- [Winkelkemper, 1983] Winkelkemper, H.E. 1983. The graph of a foliation. *Annals of Global Analysis and Geometry*, **1**(3), 51–75.
- [Xu, 1992] Xu, P. 1992. Symplectic groupoids of reduced Poisson spaces. *Comptes Rendus de l'Académie des Sciences, Série 1*, **314**(6), 457–461.

Index

- α -fibre, 80
- absolutely continuous map, 87
 - with bounded derivative, 88
- AC map, *see* absolutely continuous map
- ACB map, *see* absolutely continuous map
 - with bounded derivative
- admissible controls, **32**, 101
- admissible path, 14
 - integration, 21
 - reparametrisation, 27, 29
- AL algebroid, *see* almost Lie algebroid
- algebroid bracekt, 11
- algebroid homotopic paths, 24
 - relative to a pair of morphisms, 24
- algebroid homotopy, 17, **18**, 96
 - class
 - relative, 24
 - class, 24, 29
 - integration, 21
 - uniqueness, 19
 - with fixed end-points, 18
- almost Lie algebroid, 12
 - characterisation, 26
- Atiyah algebroid, 12, **84**
- Atiyah sequence, 84
- Carathéodory solution, 93
 - existence and uniqueness, 94
 - regularity, 94, 95
- complete lift, 13
- composable admissible paths, 14
- composition of admissible paths, 14
- composition of algebroid homotopies, 19
- cone, 103
- control, 31
- control system, 101
 - control system on AL algebroid, 31
 - non-autonomous, 41
- controlled pair, **32**, 101
 - extended, 41
- cost function, 32
 - non-autonomous, 41
- E-homotopy, *see* algebroid homotopy
- E-path, *see* admissible path
- Euler–Lagrange equations, 46
- Euler–Poincaré equations, 46
- exterior differential, 13
- final-point E-homotopy, 18
- gauge groupoid, 80
- Gronwall Inequality, 95
- Hamiltonian vector field, 13
- Hammel equations, 46
- inclusion map, 79
- initial-point E-homotopy, 18
- integrable Lie algebroid morphism, 81
- Lebesgue measure, 87
- Lebesgue theorem, 87
- Leibniz rule, 11
- Lie algebroid, 12
 - integrable, 81
 - of a Lie groupoid, 12, **81**
- Lie functor, 81
- Lie groupoid, 79
 - α -simply connected, 80
- Lie theory for Lie algebroids, 81
- local coordinates, 11
- Luzin theorem, 87

- measurable map, 87
 - bounded, 87
- measurable solution of ODE, *see* Carathéodory solution
- morphism of Lie groupoids, 79
- morphism of skew-algebroids, 14
- OCP, *see* optimal control problem
- optimal control problem, **32**, 101
 - non-autonomous, 41
 - relative to the pair of morphisms, 32
 - with general boundary conditions, 101
- pair groupoid, 80
- parallel transport, 37
- PMP, *see* Pontryagin maximum principle
- Pontryagin maximum principle, **39**, 40, 102
 - for general boundary conditions, **39**, 40, 102
 - for principal bundles, 43
 - non-autonomous, 42
- principal bundle, 83
- principal connection, 85
- product of skew-algebroids, 15
- reduced Euler-Lagrange equations, 46
- reduction map, 81
- regular point, 87
- right translation, 80
- right-invariant vector field, 80
- separation of convex sets, 103
 - cones, 103
 - strict, 103
- separation theorem, 103
- skew-algebroid, 11
 - as a linear bi-vector, 12
 - local description, 12
- source map, 79
- Stokes theorem, 20
- tangent algebroid, 12, **81**
- target map, 79
- total cost, **32**, 101
- trace, 96
- trajectory of a control system, **32**, 101
 - non-autonomous, 41
- uniformly regular map, 88
 - properties, 89
- W-solution, *see* weak solution
- weak solution, 96
 - with well defined trace, *see* WT-solution
- WT-solution, 96
 - existence, 96
 - uniqueness, 98